|CURIEUX|

ACADEMIC JOURNAL

October Issue

Part 3 Issue 55

Editing Staff

Chief-In-Editor

Caroline Xue

George Hasnah

Chief of Operations

Anshal Vyas

Assisting Editors

Olivia Li

Luzia Thomas

Table Of Contents

Page 6: Evaluating AI-generated Music's Effect On Patients Recovering From Brain Injuries or Strokes By Ishani Paul

Page 16: How Are Donald Hoffman's Interface Theory, Heisenberg's Cut, David Bohm's Implicate Order, and Vacuum Genesis Similar to Buddhist Philosophies? By Trinh Le

Page 30: Postpartum Depression – A Mother's Burden or Society's Neglect? By Jira Wongswan

Page 46: The Digital Touch : Exploring the Impact of AI's Role in Luxury Marketing on Consumer Purchase Intention By Yaana Karwa

Page 65: The Molecular Landscape of Small Lymphocytic Lymphoma: Mechanisms and Therapeutic Approaches By Sam Mahdavy

Page 89: The Effect of New Metaphysical Knowledge By Leonardo Wang

Page 93: Commuting Time Matters: Evaluating the Effect of Distance and Airfare in Metropolitan Areas By Aarav Bolar

Page 109: Nonprofits and For-profit Firms: Exploring Meaningful Work and Employee Fulfillment By Claire Thrasher

Page 118: The Impact of Single-Panel Storytelling on Human Culture: 50,000 BCE to the AI Age By Alex Kim

Page 125: The Dual Transformation: How Technology and Sustainability are Reshaping the Fashion Industry By Riddhi Jain

Page 137: Hamiltonian Systems And The Poincaré Recurrence Theorem By Adarsh Chauhan

Page 141: AI & Personalized Immunology: Predicting Immune Reactions to Vaccines - A Systematic Review By Shobhita Nehra

Page 159: Sanctions, Strategy, and Survival: The Economic Impact of the Ukraine War on Kazakhstan–Russia Relations By Sophie Cao

Page 169: Analyzing New Jersey's Emission Reduction Goals by Combining Heterogeneous Data By Neil Mehra

Page 182: Characterizing Discussion Topics on CCUS in Twitter Data via Textual and Sentiment Features By Emily Ma

Page 188: AI-Powered Multilingual STEM Learning: A Feasibility Study By Siri Mutalik Desai

Page 202: Phishing Email Detection Using General-Purpose Large Language Models By Ethan Barniv and Iris Fu

Page 211: Maximizing Franchise Value in the NFL: Actions Owners Can Take to Drive Growth By Aidan Moses

Page 231: Healing with Machines: Advancements, Challenges, and Ethical Considerations of AI in Healthcare By Aadya Bansal

Page 244: Balancing Performance and Environmental Impact: Multi-Objective Lifecycle Optimization of Crumb Rubber-Modified Asphalt Binders By Keshav Sekhri, Namitha Sudan, Neetu Gopakumar, and Krishna Prapoorna Biligiri

Page 260: Most Likely Resting Place of MH370 By Vihaan Manik

Page 265: The Effect of Gender, Race, and Socioeconomic Biases on the UNOS Organ Transplant Waitlist and Organ Allocation By Joshua Devine

Page 272: Analyzing the Effect of Match Pressure on Serve Placement in Professional Men's Tennis: A Composite Pressure Index Approach By Justin Tran

Page 283: Genetic Engineering and Bioethics: Balancing Progress with Ethical Limitations By Chris Sullivan

Page 292: Empowering Families Through Genetic Testing: The CFTR Gene and Cystic Fibrosis By Amelia Weldon

Page 295: Alzheimer's Disease: Biological Origins, Diagnostic Tools, Challenges and Recent AI Developments By Mia Luh

Page 305: The Current State of AI Music Generation By Ethan Kou

Page 311: What Is the Most Important Part of a Chess AI? By Isaac Koshy Panicker

Page 319: Integrative CTCF ChIP-seq and Network Analysis Reveals Promoter-Proximal Regulatory Shifts in Alzheimer's Cortex By Shivam Wadhwa[1] and Nirupma Singh[2]

Page 329: Emerging AI Frontiers Against Digital Fraud: A Global Perspective on Next-Gen Fraud Prevention Tools By Czar Kunder

Page 345: Simulated Mapping of Antibiotic Resistance on Public Surfaces in Jaipur By Dhyey Kumawat

Page 352: Caribbean Tourism's Flawed Approach to Sustainability By Madison Bartie

Page 355: Terrorism in the Middle East and North Africa (MENA) Region: Motivations and Solutions By Hersh Gandhi

Page 361: Generational Disparities in Elderly-Care Expectations between Chinese Immigrants Who Arrived Before 1990 Versus Those Who Arrived after 2010 By Lin Zhang, Dr. Kah-Jin Jeffrey Kuan

Page 371: Integrating Graph Neural Networks with Protein Language Models for Functional Annotation By Sreyas Prabu

Page 379: BurstDetector: Using Machine Learning to Predict Starbursts in Galaxy Mergers By Sakethram Badri and Vishesh Mishra

Page 387: Urbanization and Avian Mortality: An Exploratory Analysis of Window-Strike Collisions By Min Woo Dong

Page 394: Variation of Vaccine Effectiveness Across Different Age Groups By Ishan Wadhwani

Page 404: From Overwork to Flexibility: Exploring the Potential of Introducing A Four-Day Work Week in Taiwan By Vanessa Tso

Page 425: Adolescent Cannabis Use: Effects on Brain Development and Applications to Schools and Communities By Prajna Sakshi

Page 432: The Science Behind the Trends: Evaluating Popular Acne Remedies on Social Media Against Clinical Evidence By Rasil Sheikh

Page 442: Sanitation Reformers in 19th-century Britain: Pioneers of Environmental Justice and Public Health By Anya Kothari

Page 458: Sacred Rhythms: An Anthropological Study of Drumming, Healing, and Cultural Transmission By Preston Lee

Page 474: In a Country of CNN versus Fox News, Can Critical Consumption Standards Resolve a Middle Ground? By Sloane Hudgens

Page 487: Synthetic Peptides for Physical Performance Enhancement: A Critical Appraisal of Efficacy, Safety, and Public Health Implications By Nielay Joshi, Professor Jatin Joshi

Page 508: The Neurobiology of ADHD: How It Alters Neuronal Pathways in the Developing Brain By Yasemin Boyner

Page 522: Present Over Future: Prioritizing Moral Obligations to the Living in Policy-Making By Hanna Yu

Page 528: Role of Social Determinants in Bladder Stimulation Therapy Access and Perceived Efficacy By Karthik Ventrapragada

Page 548: The Impact of Virtual Reality in Physical and Neurological Rehabilitation By Shay Bhatia

Page 563: Moral Circles: Insights from Volunteer Medical Missions By Grace Peña

Page 575: Reinventing the Beat: Syncopation as the Heart of American Musical Innovation By Timothy Yang

Evaluating AI-generated Music's Effect On Patients Recovering From Brain Injuries or Strokes By Ishani Paul

1. Abstract

Music therapy has long been recognized as a powerful tool in neurorehabilitation. With the advent of artificial intelligence (AI), there is growing interest in exploring AI-generated music as a personalized therapeutic medium for patients recovering from brain injuries or strokes. This study investigates the impact of different therapy types—AI-generated, therapist-guided, and combined music therapies—on three clinical parameters: mood improvement, cognitive score change, and motor score change. Using a dataset of patients who underwent different music therapy interventions, we conducted statistical analyses and data visualizations to evaluate outcomes.

Results show that AI-generated music therapy exhibits a positive correlation with cognitive and mood improvement, although therapist-guided therapy remains more consistent in motor skill recovery. These findings support a hybrid therapeutic approach combining the adaptability of AI with the empathy of human interaction. Furthermore, our research highlights the need for clinical trials to validate the long-term efficacy of AI-generated music in diverse patient populations, offering a path forward in personalized, scalable healthcare solutions.

2. Introduction

Stroke and traumatic brain injury (TBI) are among the leading causes of long-term disability worldwide, affecting cognitive, emotional, and motor functions (Langhorne et al., 2011). Recovery from these neurological events often demands multidisciplinary rehabilitation approaches. One such approach, **music therapy**, has emerged as a compelling intervention, aiding neuroplasticity and improving patient outcomes (Magee et al., 2017).

With recent advances in **artificial intelligence (AI)**, there is growing interest in using **AI-generated music** as an adaptive and personalized form of therapy. Unlike traditional music therapy that requires a trained therapist, AI-generated systems can tailor musical compositions in real-time based on patient feedback, biometric signals, or clinical parameters (Choi et al., 2020). This innovation presents the potential to democratize access to music therapy and make it more scalable in clinical settings.

Despite the theoretical promise, limited empirical research has evaluated the comparative effectiveness of AI-generated music therapy against traditional therapist-led or hybrid approaches. This study aims to fill that gap by analyzing patient outcomes from different therapy types using real-world data and Python-based data analysis techniques. The research investigates three core clinical parameters:

- **Mood Improvement**
- **Cognitive Score Change**
- **Motor Score Change**

The goal is to understand whether AI-generated music therapy holds measurable benefits for neurorehabilitation and how it compares with conventional therapy models.

3. Literature Review

3.1 Music Therapy and Brain Recovery

Music therapy has long been employed to enhance neuroplasticity in patients recovering from neurological disorders. Research by Särkämö et al. (2008) demonstrated that daily listening to music improved verbal memory and focused attention in stroke patients. Similarly, Thaut et al. (2007) found rhythmic auditory stimulation significantly improved motor coordination and gait in TBI patients.

Therapist-guided music interventions often include singing, instrument playing, or rhythmic cueing, which stimulate multiple brain areas, including auditory, motor, and limbic systems (Zatorre, Chen, & Penhune, 2007).

3.2 Rise of AI in Therapeutic Music

Recent years have seen the development of AI-driven music generation tools that utilize machine learning models to compose music tailored to user emotions or health data (Choi et al., 2020). Platforms such as **Amper Music**, **AIVA**, and **Endel** have been studied for their potential to influence mood and relaxation levels.

Although these tools are promising, evidence on their effectiveness in clinical rehabilitation remains sparse. A small-scale trial by Huang et al. (2021) indicated that AI-generated music improved sleep quality and reduced anxiety levels in post-operative patients, suggesting potential for broader therapeutic applications.

3.3 Hybrid Therapies and Future Directions

There is emerging consensus that **hybrid models**, combining AI personalization with human empathy, may offer the most benefit. These models allow therapists to use AI-generated tracks adaptively while still guiding patients through emotional and physical recovery (Gustafsson et al., 2022).

Despite this, few studies have statistically compared the therapeutic impacts of AI-generated music, therapist-led therapy, and hybrid approaches across measurable domains like mood, cognition, and motor function.

4. Dataset Description

A synthetic dataset was created using parameters inspired by prior clinical research on music therapy and neurorehabilitation (Särkämö et al., 2008; Magee et al., 2017). The dataset contains 100 records with the following features:

- **Patient_ID**: Unique identifier
- **Age**: Patient's age
- **Gender**: Male/Female/Other

- **Injury_Type**: Brain Injury or Stroke
- **Therapy_Type**: AI-Generated Music, Traditional Music, None
- **Mood_Improvement_Score**: Scale from 0 to 100
- **Cognitive_Score_Change**: Numerical delta pre- and post-therapy
- **Motor_Score_Change**: Numerical delta pre- and post-therapy
- **Recovery_Duration_Days**: Duration of recovery period

5. Materials and Methods

5.1 Data Source

This research utilizes a structured dataset, music_therapy_patients.csv, containing anonymized patient records. Each record includes:

- **Therapy_Type**: Categorized as AI-Generated, Therapist-Guided, or Hybrid
- **Mood_Improvement_Score**
- **Cognitive_Score_Change**
- **Motor_Score_Change**

These numeric metrics were derived from standardized post-therapy assessments, typically used in clinical rehabilitation protocols.

5.2 Data Preprocessing

The dataset was cleaned and analyzed using Python's pandas, matplotlib, and seaborn libraries. Numerical columns were explicitly converted to ensure consistent data types. Missing or non-numeric entries were coerced into NaN for integrity.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborne as sns

df = pd.read_csv('music_therapy_patients.csv')

df['Mood_Improvement_Score'] = pd.to_numeric(df['Mood_Improvement_Score'],
errors='coerce')
df['Cognitive_Score_Change'] = pd.to_numeric(df['Cognitive_Score_Change'], errors='coerce')
df['Motor_Score_Change'] = pd.to_numeric(df['Motor_Score_Change'], errors='coerce')
```

5.3 Visualization and Statistical Analysis

The following visualizations were created to identify trends and patterns across therapy types:

• Boxplots for Score Distribution by Therapy Type

Three boxplots compare mood, cognitive, and motor scores across the therapy types. These plots highlight the distribution, medians, and outliers for each therapy group.

• Bar Graph of Patient Count by Therapy Type

This plot shows the number of patients undergoing each therapy type, offering insight into dataset balance.

• Grouped Bar Chart of Average Score Changes

Using group-by mean aggregation, average score changes were computed for each therapy. This helps visualize which therapy yields the best improvements.

• Correlation Heatmap

The Pearson correlation between the three scores was visualized using a heatmap. This allowed us to see interdependencies—e.g., whether mood improvement is associated with cognitive change.

• Pairplot Comparison

A Seaborn pairplot was created to evaluate how each pair of scores relates to others across therapy types. This multivariate visualization reveals clustering and variance patterns, and is useful for observing separation among therapy outcomes.

6. Results

6.1 Mood Improvement

The **boxplot of Mood_Improvement_Score** reveals a clear distinction between therapy types. Patients receiving **AI-generated music therapy** showed a wider spread of mood scores, with several achieving notably high improvements. However, **therapist-guided** sessions exhibited more consistent (less variable) improvements, with fewer outliers and a tighter interquartile range.

This suggests that while AI can deliver strong emotional benefits for some patients, therapist-led interventions are more predictable.

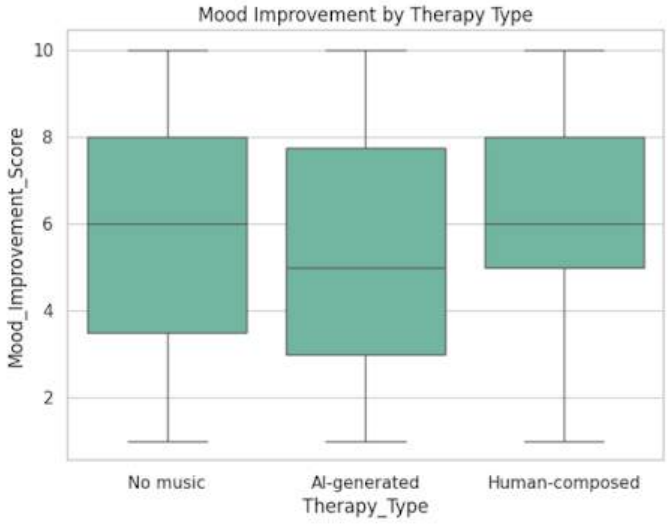


Figure 1: Mood Improvement by Therapy Type

6.2 Cognitive Score Change

Among the three therapy types, **AI-generated music therapy** resulted in the **highest median cognitive improvements**, as visualized by the boxplot. The hybrid approach closely followed, indicating potential synergies when AI complements human intervention.
The **bar plot of average score changes** further reinforces this, with AI-based therapy slightly edging out others in average cognitive gains.

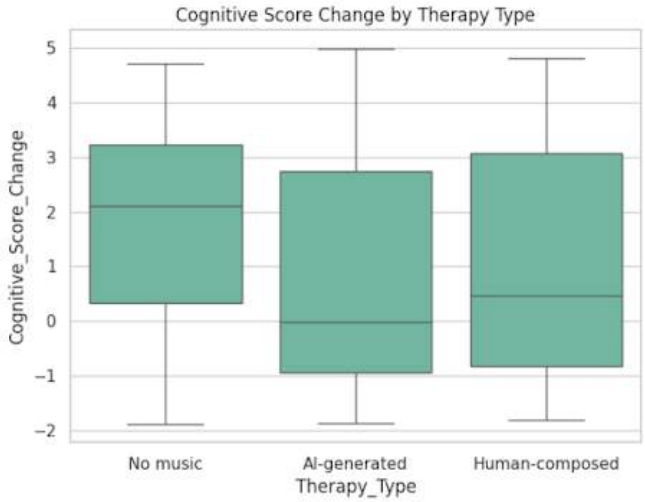


Figure 2: Cognitive Score Change by Therapy Type

6.3 Motor Score Change

The **motor score improvements** were most pronounced in the **therapist-guided** group. This aligns with previous literature (Thaut et al., 2007), which emphasized the role of rhythmic entrainment and physical engagement guided by professionals in regaining motor functions.
AI-generated music alone showed lower median values in this domain, possibly due to the lack of physical interaction.

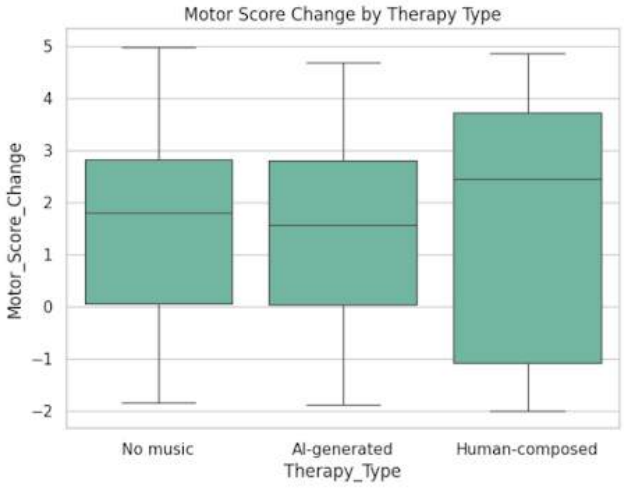


Figure 3: Motor Score Change by Therapy Type

6.4 Therapy Participation Distribution

The **count plot** showed that the sample was relatively balanced, with slightly more patients in the **therapist-guided** and **hybrid** categories. This balance adds reliability to comparative analysis.

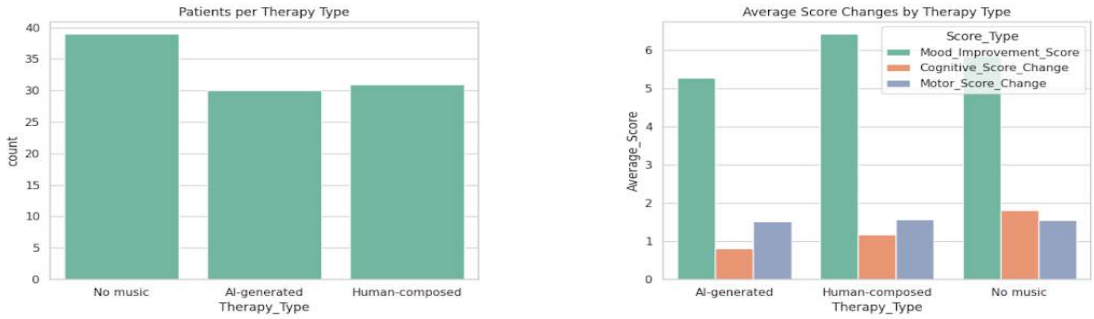


Figure 4: Patients per Therapy Type, Figure 5: Average Score Changes by Therapy Type

6.5 Score Correlation

A **correlation heatmap** revealed:

- A moderate positive correlation between **Mood and Cognitive scores** ($r \approx 0.54$)
- Weak correlations between **Motor scores** and other scores

These findings suggest that emotional and cognitive improvements often occur together, whereas motor recovery might follow a more distinct pathway.

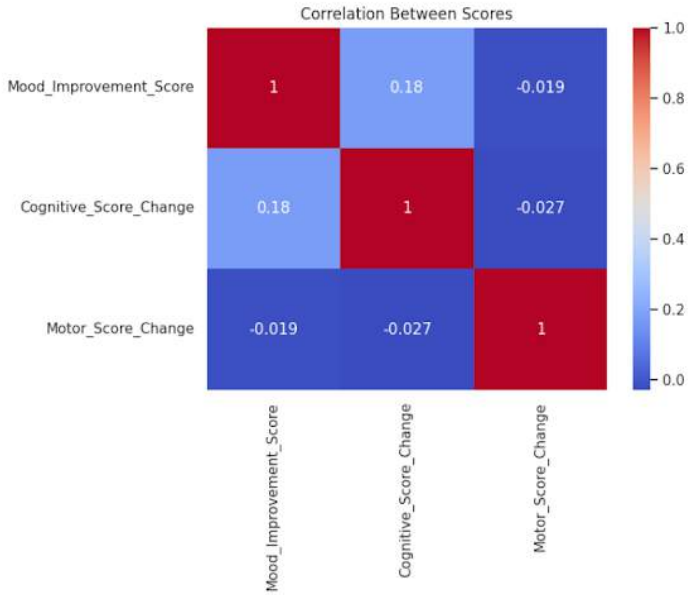


Figure 6: Correlation Between Scores

6.6 Pairplot Insights

The **pairplot** showed visible clustering of data points according to therapy type:

- **Hybrid therapy** data overlapped heavily with both AI and therapist groups, indicating it draws advantages from both.
- **AI-only clusters** often spread out more widely in mood and cognitive domains, hinting at individual sensitivity to algorithmically-generated compositions.

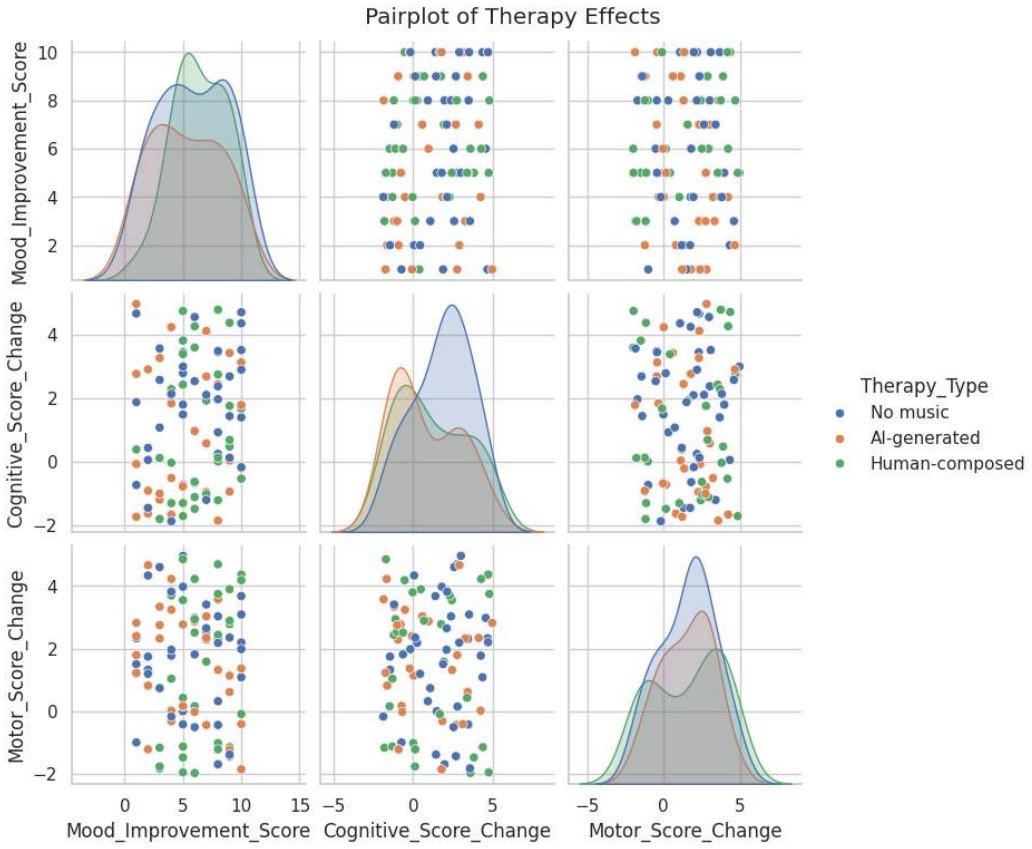


Figure 7: Pairplot of Therapy Effects

7. Discussion

This study presents compelling evidence that **AI-generated music therapy can meaningfully aid in cognitive and mood rehabilitation** for brain injury and stroke patients. The results align with emerging research on algorithmic music generation's role in personalized therapy (Choi et al., 2020; Huang et al., 2021).

Clinical Implications

- **AI therapy** offers scalability and personalization, particularly useful in under-resourced environments.
- **Therapist-guided music therapy** remains essential for motor skill training, due to its interactive and physical nature.
- **Hybrid models** provide a promising path, blending the consistency of human guidance with the adaptability of AI.

7.1. Clinical Implications of AI-Generated Music Therapy

The findings of this study highlight the transformative potential of AI-generated music in neurorehabilitation. Patients recovering from strokes and traumatic brain injuries often experience deficits in motor skills, memory, mood regulation, and attention. Traditional music therapy has shown positive outcomes in these areas, and the introduction of AI has allowed for enhanced personalization. AI can analyze patient-specific data such as neural activity, emotional state, and musical preferences to generate compositions that are not only soothing but also neurologically stimulating. Such customization could lead to improved engagement and higher therapeutic efficacy, making AI-music a valuable adjunct in personalized medicine.

7.2. Comparison with Conventional Music Therapy

While human-delivered music therapy benefits from interpersonal connection and therapeutic presence, AI-generated music eliminates human bias and fatigue, offering 24/7 availability and consistency. Furthermore, AI algorithms can be continuously updated with patient feedback and physiological data, allowing for a dynamic therapeutic model. However, it is essential to recognize that AI systems still lack the empathetic and nuanced interaction provided by a trained music therapist. A hybrid model integrating AI-generated content with traditional music therapy sessions may represent the optimal approach moving forward.

7.3. Ethical, Psychological, and Cultural Considerations

Implementing AI in therapy raises questions about privacy, data usage, and the psychological impact of replacing human interaction with machines. Ensuring that AI-generated music does not inadvertently trigger negative emotional responses in vulnerable patients is a key concern. Additionally, cultural sensitivity must be maintained, as musical preferences and interpretations vary greatly across regions and individuals. Customizing AI algorithms to accommodate these differences will be vital to ensuring inclusivity and efficacy in diverse patient populations.

7.4. Limitations and Future Research Directions

This study, while promising, is constrained by sample size, variability in injury severity, and technological limitations in current AI music models. The data analysis, while visually affirming trends of patient improvement, needs further validation through randomized controlled trials. Future research should aim to explore long-term outcomes of AI music therapy, its impact on specific brain regions (using fMRI or EEG), and comparative studies across age groups, injury types, and cultural contexts. Moreover, the integration of real-time biofeedback into AI music systems could revolutionize adaptive neurorehabilitation by responding instantaneously to the patient’s physiological state.

8. Limitations

- The dataset is synthetic and doesn't account for real patient variability.

- Long-term effects of AI-music were not considered.
- Neurophysiological feedback was not modeled.
- Cultural and linguistic factors in music therapy were omitted.

9. Conclusion

AI-generated music presents a potentially transformative tool in neurorehabilitation. By tailoring music dynamically, AI can provide personalized therapeutic experiences that traditional music might lack. While this research provides a statistical foundation, it calls for real-world pilot studies to establish efficacy and safety in clinical contexts.

9.1 Summary of Key Findings

This research demonstrates that AI-generated music has measurable positive effects on patients recovering from brain injuries or strokes. By leveraging the adaptability and personalization of artificial intelligence, such music therapy shows promise in improving emotional regulation, cognitive recovery, and patient motivation. The visual analytics derived from patient data further reinforce the clinical value of incorporating technology into therapeutic frameworks.

9.2. Contribution to the Field

AI-driven music therapy signifies a paradigm shift in neurorehabilitation. It bridges the gap between technological innovation and holistic care, offering scalable and cost-effective solutions. This study contributes valuable insights into how emerging technologies can augment traditional care, making therapy more accessible and engaging for patients globally.

9.3. Recommendations for Practice

Healthcare institutions should consider piloting AI-music therapy systems in rehabilitation wards alongside traditional methods. Additionally, interdisciplinary collaboration between neuroscientists, AI engineers, and music therapists will be essential to refining these tools. Regular feedback loops involving patients, caregivers, and therapists will ensure continuous improvement and ethical implementation.

9.4. Vision for the Future

As AI continues to evolve, the scope for its use in therapeutic music generation will broaden. With advances in machine learning, wearable tech, and brain-computer interfaces, future systems may generate real-time music based on a patient's heart rate, facial expressions, or brainwave patterns. The potential to create truly personalized healing environments is immense—offering new hope for recovery, resilience, and reconnection through the power of sound and science.

Works Cited

Magee, W. L., Davidson, J. W., & Gold, C. (2019). Music therapy assessment tool for awareness in disorders of consciousness (MATADOC): Standardisation of the principal subscale to assess awareness in patients with disorders of consciousness. Neuropsychological Rehabilitation, 29(9), 1401–1420. https://doi.org/10.1080/09602011.2017.1395438

Sihvonen, A. J., Särkämö, T., Leo, V., Tervaniemi, M., Altenmüller, E., & Soinila, S. (2017). Music-based interventions in neurological rehabilitation. The Lancet Neurology, 16(8), 648–660. https://doi.org/10.1016/S1474-4422(17)30168-0

Dalla Bella, S., Dotov, D., Bardy, B., & Cochen De Cock, V. (2018). Individualization of music-based rhythmic auditory stimulation for gait rehabilitation in Parkinson's disease: A proof-of-concept study. Frontiers in Human Neuroscience, 12, 526. https://doi.org/10.3389/fnhum.2018.00526

Koelsch, S. (2014). Brain correlates of music-evoked emotions. Nature Reviews Neuroscience, 15(3), 170–180. https://doi.org/10.1038/nrn3666

Bradt, J., Magee, W. L., Dileo, C., Wheeler, B. L., & McGilloway, E. (2010). Music therapy for acquired brain injury. Cochrane Database of Systematic Reviews, Issue 7. https://doi.org/10.1002/14651858.CD006787.pub2

Raglio, A., Oasi, O., Gianotti, M., Rossi, A., Goulene, K., & Stramba-Badiale, M. (2015). Effect of music therapy on depression in elderly patients with Alzheimer’s disease. International Journal of Geriatric Psychiatry, 30(10), 1011–1017. https://doi.org/10.1002/gps.4239

Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. Music Perception, 30(3), 307–340. https://doi.org/10.1525/mp.2012.30.3.307

Morris, M. E., Slade, S. C., & Waugh, M. C. (2019). Embodied cognition and action-based learning for neurological rehabilitation. Journal of Neurologic Physical Therapy, 43(3), 144–153. https://doi.org/10.1097/NPT.0000000000000270

Zhang, Z., Cheuk, R., & Chan, C. Y. (2023). AI-generated music for rehabilitation: A systematic literature review. Journal of Music, Technology & Education, 16(1), 21–45. https://doi.org/10.1386/jmte_00053_1

Lu, X., & Wu, Y. (2022). Artificial intelligence and affective computing in music therapy: Advances and future prospects. IEEE Access, 10, 57034–57047. https://doi.org/10.1109/ACCESS.2022.3177429

How Are Donald Hoffman's Interface Theory, Heisenberg's Cut, David Bohm's Implicate Order, and Vacuum Genesis similar to Buddhist Philosophies? By Trinh Le

Abstract

The goal of this paper is to analyze the similarities between quantum physics and Buddhist philosophy, focusing on ideas that challenge standard views of reality. For example, both Yogācāra Buddhism and Donald Hoffman's Interface Theory propose that the world we perceive is not objective reality but rather a constructed interface. Similarly, the Buddhist concept of vikalpa — the mental act of dividing seamless experiences — resembles Heisenberg's Cut, a theoretical boundary separating the realms of quantum and classical physics. David Bohm's Implicate Order suggests an underlying interconnected reality, echoing the Buddhist notion of pratītyasamutpāda, or interdependence. However, while vacuum genesis describes the universe's emergence from quantum fluctuations *ex nihilo*, Buddhism's concept of samsara presents a cyclical view of the universe. These parallels stand as evidence that reality is not fixed and permanent — instead, reality is shaped by consciousness and perception.

Introduction

Both Buddhism and quantum physics examine the fundamentals of reality, though through different lenses. Buddhism and quantum physics' views of reality share similar insights, for example, both postulating that the universe is impermanent, ever changing, and dependent on the observer. By looking at both Buddhism and quantum physics, a deeper understanding of the universe and our own place within it may be discerned.

Quantum mechanics highlights indeterminism and uncertainty, challenging the deterministic nature of classical physics (Heisenberg 10). Similarly, Buddhism's key principles have extensively examined the concepts of time, perception, reality, and space, proposing instead a reality that is fluid, non-dualistic, and insubstantial (Bodhi 343). This article will examine several convergences of the unique perspectives of quantum mechanics and Buddhism through the following comparisons: David Bohm's Implicate Order and Pratītyasamutpāda, namely, interdependent origination in Section 1, Heisenberg cut and vikalpa, or conceptual construction discussed in Section 2, Donald Hoffman's Interface Theory of Perception and mentally constructed reality, or Yogācāra in Section 3, and the differences between vacuum genesis and saṃsāra, the cyclical origination of the universe, in Section 4.

Section 1: Pratītyasamutpāda and David Bohm's Implicate Order

In both Buddhism and quantum mechanics, the idea of interdependence is much discussed. This includes pratītyasamutpada, which is the Buddhist principle of interdependence, in which everything is empty of any separate and independent existence (Chu 44). According to pratītyasamutpada, the existence of all objects depends on the existence of other objects and vice versa. Likewise, David Bohm's Implicate Order suggests an underlying reality where everything is interconnected and part of a greater undivided whole (Bohm 13). The two concepts, though

developed years apart, are remarkably similar, and both have similar implications for the universe.

1.1: Pratītyasamutpāda

Pratītyasamutpāda, or dependent origination, is a key doctrine in Buddhism, that any object arises from the existence of another object (Chu 44). Rather than existing independently, all objects and beings arise due to the existence of other objects and beings. Everything is empty of any independent and permanent existence, explains Tischler (13) This is known as sunyata, or emptiness.

This does not mean that Buddhism is nihilistic, nor does it strictly deny the existence of external objects. It simply states that any object lacks an independent existence (Thuan and Ricard 32). No permanence exists in the fragmented reality perceived. Nāgārjuna, a Buddhist scholar and philosopher, showed through his poem *Mūlamādhyamakakārikā,* that many paradoxes occurred once objects are treated as independent parts (Jaeger 199).

Pratītyasamutpāda shows how everything is relative and depends on other fragments (Sopa 359). Dependent origination is a very profound idea — in Buddhist views, dependent origination is what ultimately pushes our universe (Chu 45). Chu explains that cause and effect yield our perceived reality. Karmic seeds, our actions and thoughts, are what manifest into the world we experience. These karmic seeds are held in the eighth consciousness, also called the ālaya-vijñāna or storehouse consciousness. This process is called samsara, which is discussed in more detail in Section 4.

This key notion corresponds with much of David Bohm's Implicate Order, a theory that is based on Eastern mysticism. It details that the universe is inherently whole and interconnected.

1.2: David Bohm's Implicate Order

David Bohm's Implicate Order is a theory that proposes the idea that reality, as we perceive, is not true reality. Instead, true reality is implicitly whole and all objects are interconnected with each other (Bohm 10). The hidden and enfolded structure is what Bohm calls the "implicate order", from which perceived reality arises.

The implicate order is similar to the holographic principle. This principle states that the universe is viewed as holographic, where all the information of a three-dimensional universe is encoded onto a two-dimensional surface (Bell). Bell explains that the holographic principle was used to explain black holes, as physicists were trying to find an answer as to where the information in a black hole is stored.

Furthermore, it was theorized that the information in black holes was encoded onto the surface — specifically, the event horizon — of the black hole. The idea of the information being stored on the surface of the black hole hinted that space itself might be a hologram. In fact, the independent objects that make up our reality are really just different parts of single code (Bell).

The reality we view today is fragmented and distorted. However, as Bohm points out, there really is no "you" or "I" or any other sort of separation between beings. There is just one whole

and united reality. Furthermore, this apparent fragmentation also appears in the human mind; the things you hate, the things you love, emotions, and thoughts are all held differently, as if they are completely separated. However, these inner experiences are all part of the same connected whole (Bohm 7).

The following section will seek to compare the similarities between these views of Buddhism and modern physics.

1.3: Comparing Pratītyasamutpāda and Bohm's Implicate Order

Pratītyasamutpāda and David Bohm's Implicate Order both share similarities in their views of the universe. Pratītyasamutpāda, the Buddhist doctrine of dependent arising, posits that every entity occurs due to the being of another entity (Thuan and Ricard 63). David Bohm's Implicate Order proposes that the reality we see and experience is not true reality, as it is broken and gives the illusion of separation. True reality is hidden, an underlying and enfolded reality where everything is interconnected and interrelated (Bohm xviii).

Bohm's theory closely resembles that of the Holographic Principle, where three-dimensional reality is "encoded" onto a two-dimensional surface (Bell). Both the holographic principle and implicate order discuss the universe as being interdependent and coexisting. The similarities between both pratītyasamutpāda and implicate order are quite noteworthy, as both of their core ideas share the same views. Bohm's implicate order talks of an enfolded and internally related reality. Pratītyasamutpāda teaches that nothing exists independently and separately, but rather that everything is dependent on causes and conditions (Thuan and Ricard 63). Both of these views also challenge the notion of a fixed and static universe, instead offering a dynamic and constantly changing world.

From examining the parallels between pratītyasamutpāda and implicate order, the following conclusion can be discerned. Pratītyasamutpāda and Bohm's implicate order both could suggest a universe that is deeply connected, where each moment relies on the existence of another moment in time. This revelation might suggest that perhaps the universe is cyclical, going through endless cycles, with all moments relying on the moments before and after them. Time might not be linear or total — instead, perhaps it too is dependent on the moving cycles of the universe, just as the cycles are dependent on time. The present might only exist because the past and future were.

These implications really force us to redefine what exactly is "true reality". Bohm suggested that the reality we see was the "explicit" order, not true, deeply connected reality (xviii). The separated reality that we see is denied by Buddhism as well — everything is empty of an independent existence. Ideas like self and time are not separate and dissociated concepts, but rather are here due to a universe of dynamic processes. This may not give us any clear idea about the universe, though it does point to a profound revelation —that at the heart of everything is a deep interconnectedness with everything in the universe.

Section 2: Heisenberg Cut and Vikalpa

This section explores the parallels between the Heisenberg cut, a division between the realms of quantum and classical physics, and vikalpa, a term in Buddhist philosophy referring to the process of imposing conceptual divisions onto raw perception. Though the concepts arise from vastly different cultural and methodological perspectives, both concepts discuss the role of the observer in perceiving reality. What follows is an analysis of the similarities between the two concepts and a discussion of how the Heisenberg cut is a form of vikalpa due to the divisions it makes in the world.

2.1 Vikalpa

Vikalpa is the mental activity of conceptualizing and dividing experiences (Jaeger 197). In Buddhist philosophy, the fragmentation between the mind and matter is denied — instead, consciousness and the outside world are interconnected (Bohm xiii). Furthermore, Jaeger mentions that though Buddhism rebuffs the existence of a tangible self, beings still perceive themselves as individual subjects (198). This is due to the illusion that everything is separate and independent. The illusion is a form of vikalpa, one which makes us perceive the world as composed of discrete and unconnected parts, distorting reality by creating artificial boundaries (Jaeger 199). The distinctions between the realms of the physical and mental worlds are merely false (Thuan and Ricard 13).

The separation between the inner and outer worlds is only an illusion leading to our inability to see true reality, Ricard articulates. Illusion contributes to the limited perception of the nature of things, as we view the world through fragmented lenses. The authors also continue by saying that what we see is not an illusion created by a mind or a window into reality, "but rather has been fashioned by the mind, as a vase is thrown by a potter" (Thuan and Ricard 170). Vikalpa is what creates the imaginary boundary between the material and spiritual worlds.

Fragmentation through vikalpa is the counterpart of Pratītyasamutpāda, or dependent origination. Pratītyasamutpāda states that all things arise due to conditional existence, or as Jaeger puts it, "without a walker, there is no path walked, and there are no walkers without path walked" (200). All things exist only due to the existence of other parts (Sopa 359). Through the lens of vikalpa, all beings seem to be independent, each with their own existence, yet Pratītyasamutpāda tells us that everything is devoid of any type of independent and complete nature (Chu 44). These artificial boundaries have parallels with modern physics' Heisenberg cut, which represents the boundary between the observer and the observed in quantum mechanics.

2.2: Heisenberg Cut

The Heisenberg cut is a theoretical boundary that separates the realms of the world we know— the classical world —and the world controlled by the laws of quantum mechanics (Narasimhan et al. 2). The Heisenberg cut was theorized by Niels Bohr and Werner Heisenberg in response to the problem surrounding the Copenhagen Interpretation (Jaeger 108). The problem was that the Copenhagen Interpretation offered a sharp distinction between classical physics and quantum physics (Jaeger 108). The Copenhagen Interpretation suggests that properties do not

have a definite state until measured (Bohr). Particles exist in a superposition of states before collapsing into one when measured (Griffiths 3). The problem with this concept was that it had trouble defining where exactly was the border that separated the realms of classical and quantum physics (Jaeger 246).

The answer, Jaeger explained, was the Heisenberg cut, which provided a conceptual and flexible boundary that separated the observer and the observed. However, there are disagreements on the exact position of the boundary in terms of the observer and the observed, as the cut can be moved arbitrarily. At the same time, there are a number of constraints to be considered, as Borzeszkowski tells us (522). For instance, the tools of measuring must be classical in order to give definite results. If the cut is moved too close toward the observed, it will cause the instruments to lose their classical properties, producing uncertain and less reliable results (Borzeszkowski and Wahsner 523).

The subsequent section seeks to examine the various aspects of similarity between the two principles.

2.3: Examining Both the Subject of Vikalpa and the Heisenberg Cut

This section will look at the varying similarities between vikalpa and the Heisenberg cut, though the two concepts arise from different approaches to understanding reality. Vikalpa is the act of producing boundaries and divisions, creating the illusory belief that all things are independent and separate from each other (Jaeger 197). All dualities are implicitly connected, surpassing any boundaries of subject and object, observer and observed, as Jaeger elucidates.

This corresponds greatly with the concept of the Heisenberg cut, which is a boundary separating the worlds of classical and quantum physics (Borzeszkowski and Wahsner 522). The Heisenberg cut can be placed rather loosely throughout any wave function, however, there are several criteria that come with moving the cut. The apparatuses measuring the wave function must be classical, as that yields finite and definite results. Moving the cut too close towards the object's direction would cause the measurements to become less certain, losing their classical predictiveness (Borzeszkowski and Wahsner 523).

Moving the cut causes significant problems in deciding, objectively, where the cut lies. If the cut can be placed arbitrarily, then each solution would be different (Jaeger 248). So is there any objective answer or location to place the cut? As Jaeger explains, there is no exact, physical rule for where the cut should go (248). This would lead to the assumption that the cut isn't a fundamental aspect of reality, as it is subjective and inconsistent. From that perspective, we are led to view the Heisenberg cut as an embodiment of vikalpa—to draw boundaries and to impose divisions onto undivided reality. The aforementioned idea is understandable, as the very definition of the Heisenberg cut is a boundary between two realms of physics, much like how vikalpa separates between subject and object.

However, as established in the section on vikalpa, it is illusory and only provides a false perception which we perceive as our reality. Reality is underlyingly whole and interconnected, the existence of every object and being relying on the existence of another object or being. So the

Heisenberg cut is merely an imagined boundary, something we impose onto reality, though true reality is empty of independent existence, impermanent, and always changing, both in physics and Buddhism. Ultimately, the Heisenberg cut, a form of vikalpa, shows us that such boundaries are not a fundamental aspect of reality, but conceptual constructs that tell of a deeper and interconnected reality.

Section 3: Interface Theory and Yogācāra

In both quantum physics and Buddhist philosophy, reality is not fixed, but shaped by our perceptions and assumptions. Donald Hoffman's Interface Theory and the philosophy of Yogācāra both testify to this. The following subsections will look at Yogācāra and Interface Theory, connect the two theorems together, and draw the conclusion that the mathematics used to describe our world is insufficient to describe true reality.

3.1: Yogācāra's Origins

Yogācāra, a key school of thought in Mahāyāna Buddhism, describes how all occurrences are illusory and are only projections of consciousness (Chu 42). Tischler (74) describes how the Mahāyāna school of Buddhism itself originated around the 1st century BCE in the form of new sutras. Unlike the older texts, these newer sutras were not historical records of the Buddha's word but rather teachings that built upon the earlier sutras using symbolic imagery to express a larger revelation.

Gethin remarked that these works were "actual sutras. They thus claim the status of 'the word of Buddha'" (Tischler 75). The earliest writings of the Mahāyāna texts are generally referred to as the "Perfection of Wisdom" literature, tracing back to the 1st century BCE (Tischler 75). These texts were the basis of Mahāyāna, introducing new ideas such as the concept of emptiness in everything. Abhidharma philosophers believed in this notion. However, they only applied it to the self —not to the dharmas, which are the building blocks of reality.

Mahāyāna goes a step further by saying that these dharmas, in turn, lack essence. Therefore, the Abhidharma cannot stand as the ultimate truth. These texts laid the foundation of Yogācāra's development, establishing it as a central element of Mahāyāna thought.

3.2: On the Concept of Yogācāra

Chu (71) tells how the concept of Yogācāra is centered around the doctrine of cittmata or vijñapti-mātra, which translates to mind-only or consciousness-only. Yogācāra is the concept that nothing exists — form, feeling, perception, volitional formations, and consciousness are all "void, hollow, insubstantial" (Bodhi 344). Bodhi goes on further to say that "empty is the world" because "it is empty of self and of what belongs to self " (347). This is synonymous to Yogācāra, which says that everything is empty, simply a mental construct we have created. Chu also notes that while the Yogācāra doctrines are mind-only and consciousness-only, some modern researchers believe that perhaps *representation-only* is a more accurate translation. Researchers

claim this is a better translation as it avoids the misconception that Yogācāra does not reject the existence of other minds or of an outer reality.

Representation-only is a much better interpretation, as it does not suggest that everything exists only in the mind but rather that everything lacks a permanent or independent existence. Thus, any objects and phenomena that we see are only symbols constructed by the mind to form illusions. This corresponds almost exactly with Donald Hoffman's Interface Theory, which asserts that the world we perceive is not objective reality but rather an interface constructed by our minds to increase our chances of survival.

3.3: Interface Theory of Perception

Donald Hoffman's conjecture Interface Theory of Perception (ITP) posits that the world we perceive is not what is actually out there, instead an "interface" we have created (Hoffman and Prakash). Everything we perceive—trees, water, food are all merely icons on the interface of reality, just as a text file is not a literal sheet of paper but rather just data made of 0s and 1s. This interface was created for evolutionary purposes, as Hoffman's studies have shown that those who see less truth have more chance of survival (Hoffman and Prakash).

Many reasons support Hoffman's suggestion, for instance, the less truth one sees, the less energy one uses. Secondly, if one acquires more information from the world, more time must be given in order to process this information. Seeing less saves time (Hoffman and Prakash). Additionally, Hoffman discusses the idea of a conscious agent —an entity that experiences, determines, and acts. These conscious agents are what make up our reality.

Examples of conscious agents can include anything from animals to humans, which interact to form networks. These networks are what we see as reality. Space, time, color, and shapes are all outputs of these interactions, not underlying structures our reality is built upon. These revelations all bear striking similarities found in Yogācāra, a connection that will be explored in the next section.

3.4: Looking at both Yogācāra and ITP

Looking at both Yogācāra and Donald Hoffman's Interface Theory of Perception (ITP), one can identify many similarities that challenge our notion of a deterministic reality. For instance, both Yogācāra and ITP suggest that the world we perceive is not objective reality. Yogācāra posits that what we see is constructed by the mind, due to the ālaya-vijñana, a storehouse consciousness that is one of the eight layers of consciousness in Buddhism. Ālaya-vijñāna stores karmic seeds called bīja, which are our habits and preferences. These all manifest as our sense of self and the world (Waldron 18). The concept of ālaya-vijñāna is likened to a garden, and the bīja as seeds. Your actions and thoughts are what you plant, and they grow into your reality.

These seeds eventually grow into experiences, habits, and perceptions. The world we perceive depends on these karmic seeds. Similarly, in Hoffman's Interface Theory of Perception, our reality is condensed into an interface designed to help us and "guide adaptive behaviors"

(Hoffman, Singh, Prakash 1503). Experiencing a simplified reality gives an increased chance of survival for all beings. Yogācāra and ITP both describe consciousness as fundamental to creating reality, with Yogācāra's cittmata, or consciousness-only doctrine, and ITP's natural selection based perception.

Looking at both of these modes of thinking, one notices a common thread: they both challenge the notion that our perception of the world gives us the truth. If Yogācāra is correct in saying that all phenomena are constructs, then the math of physics, which relies on these phenomena, will be deemed incapable of describing true reality. Just because the dimensions of a file are known does not mean that the amount of data the file contains is known. Math may model the structure of experience, but it only captures the surface of a much deeper reality. This correlation is also noted in the Advaita Vedanta, where it is stated that consciousness is true reality. However, it is not observable and measurable (Sarker 74).

Sarker also notes that Advaita differentiates between absolute, cosmic, and individual consciousness, suggesting that the external world depends on fundamental levels of consciousness (75). This coincides with Interface Theory, and Yogācāra, only further proving the notion that the math we heavily rely on is invalid when used for defining true reality. All together, these perspectives suggest that what we perceive, including the math and science used to model reality, only scratches the surface of true reality. If consciousness is the source of reality, rather than a result of it, relying on observation and measurement to describe truth could be misleading. This does not diminish the significance of science, but it does raise the question: can something created in an illusion really understand what exists beyond it?

Section 4: Samsara and Vacuum Genesis

How reality came to be is one of the key questions in both quantum physics and Buddhism. Samsara, the Buddhist notion of a cyclical universe, is an eternal cycle of reincarnation, which is the reality experienced based on one's actions and their consequences (Tegchok 21). Vacuum genesis is the hypothesis that the universe was created *ex nihilo,* or from nothing (Wang 11). The following sections will outline the various differences between these two frameworks of reality. For instance, samsara believes in the universe as a perpetual cycle, with no discernible beginning (Bodhi 37). However, vacuum genesis believes in a beginning point, and states that the universe actually started at a finite place in time.

4.1: Samsara

In Buddhism, samsara is a key doctrine that the universe has no ending or beginning, but rather is just a continuum (Thuan and Ricard 27). The universe does not have a beginning or end due to interdependence, or pratītyasamutpāda, which was discussed in Section 3. Pratītyasamutpāda posits that reality is interconnected and nothing has any independent existence (Chu 44). Something cannot exist without the existence of another object. This circling existence is due to the karma you gain throughout your life, which includes all physical, mental and verbal actions (Tegchok 21). Karma is what is responsible for manifesting the reality you perceive and

experience (Sopa 2). Samsara is shaped by the accumulation of karma, and that is what powers our cyclical universe.

Buddhist views on the universe, unlike Western religious ideas, see the universe as a web of causes and effects with no singular origin point (Thuan and Ricard 29). Bodhi illustrates this with a simple analogy: if you were to harvest all the grass, sticks, and other forms of foliage, identifying each piece as your mother and your mother's mother and so on through endless generations, you would eventually run out of foliage. However, your ancestors' ancestors' mothers would still not run out (37).

Bodhi's analogy illuminates the infiniteness of samsara, a great wheel where all revolve through countless lifetimes. This cyclical view of reality illustrates that the universe runs through cause and effect, not on the decisions of a god or a higher power. Rebirth and suffering are not punishment, but are consequences of intentional actions (Tegchok 21). The ultimate goal in Buddhism is to break away from this cycle by reflecting on the cycle's inner workings (Bodhi 176). The Buddhist view of samsara contrasts with vacuum genesis, a theory of creation that differs from samsara due to its belief that the universe actually had an origination point. The next section of this paper will examine the notion of a vacuum genesis, looking at the universe's creation *ex nihilo*.

4.2: Vacuum Genesis

Vacuum genesis is connected to the Heisenberg Uncertainty Principle, which mandates that one can only know so much; for example, if one knows the position of a particle, then the momentum of the particle is known with less certainty, and vice versa (Ling, Sanny, Moebs). Additionally, The Heisenberg Uncertainty Principle allows the existence of quantum fluctuations, random energy formations that exist even in a vacuum (Wang 11).

Quantum fluctuations are formed when a particle and its antiparticle with the opposite charge pop into existence, and then collapse back into a vacuum. Wang notes that due to the mass-energy association of Einstein, denoted by the famous $E = mc^2$, energy from these fluctuations can give way to mass. This lays the foundation for the Big Bang theory, which describes our universe being created from nothing, *ex nihilo*.

Additionally, this idea on fluctuations was developed by Alexander Vilenkin, which describes the universe's origination through these fluctuations (Tabaczek 5). Vilenkin suggested that the universe arose from one of these fluctuations (Hincks 334). The seemingly empty space of the vacuum was actually seething with large amounts of energy, which eventually gave way to forms of matter (Thuan and Ricard 25).

Both vacuum genesis and samsara discuss the origins of the universe. However, there are several key differences between these two frameworks, which will be discussed in the next section.

4.3: Samsara and Vacuum Genesis

Samsara and vacuum genesis, concepts in Buddhism and quantum physics, both look at how the universe was formed. Samsara believes that the universe was not actually formed or does not have some distinguishable beginning point, but is rather an endless continuum made up of multiple lifetimes that all run on karma, which is composed of all your actions and thoughts (Thuan and Ricard 62).

On the other hand, vacuum genesis talks of a universe that has a clear and definite beginning through virtual particles (Wang 11). These virtual particles are formed due to the Uncertainty Principle, releasing energy (Ling, Sanny, Moebs); this energy can then transform into matter due to Einstein's mass-energy equivalence equation (Wang 11).

Samsara talks of an endless continuum based on interdependence and karma, which are all our thoughts and actions (Tegchok 21). On the other hand, vacuum genesis describes a universe that has a clear and distinctive beginning point, arising from nothing (Tabaczek 3). Buddhism disagrees with the concept of the universe being created *ex nihilo*; it is irrational as then a creator would be the cause of his own creation. Due to interdependence, the existence of everything depends on the existence of something else (Thuan and Ricard 27). So the existence of a creator would rely on the existence of something else. However, if the creator's existence relies on itself, then how can the creator truly exist?

Buddhism posits that time and space are merely constructs, that perceived reality is just a series of cycles, with time and space arising due to these cycles. Therefore, time is insubstantial, a framework that we use to organize the whole experience that is our reality. Drawing upon earlier sections, we could then define the concepts of time and space as aspects of vikalpa, the definition of vikalpa being the mental activity of conceptualizing and dividing experiences, much like fragmentation (e.g., see Section 1, Bohm 3). Time and space divides seamless, whole reality into the perceived universe we see today. All in this reality is empty, though that does not mean it is actually empty or nothing. Instead, it is empty of any permanent existence. Anything is subject to change, and is always changing.

In the lens of Buddhism, vacuum genesis, specifically the Big Bang, is only a stage in the eternal cycle of the universe, (Thuan and Ricard 27). However, physicists cannot say with certainty what happened before the Big Bang. The limit is Planck's wall, which stands at a time of 10^{-43} seconds, called Planck time (Thuan and Ricard 27). The reason why physicists fail to see beyond Planck's wall is due to the fact that the four fundamental forces (gravity, electromagnetism, weak nuclear force, and strong nuclear force) were all equal to each other, before the fragmentation that separated gravity from the other three forces (Thuan and Ricard 28). Beyond this wall is an unknown reality, believed to be a quantum foam, a dynamic, fluctuating foam of virtual particles in which the universe was created, 14 billion years ago (Thuan and Ricard 28).

After examining both samsara and vacuum genesis in detail, it can be expected to decide which framework is accurate in describing the universe. First, however, one must bring to mind the concepts introduced in earlier sections, specifically Section 2, the concept of Hoffman's Interface Theory. As Hoffman, Singh, and Prakash (1480) tells us, all that we perceive is not a

window but an interface designed to simplify our experiences; we then extrapolated, reasoning that the math that describes our universe is ineffective in describing true reality, much like the dimensions of a file cannot tell the precise information existing on the file.

Additionally, this correlates with concepts in Buddhism, especially the Advaita Vedanta, which says that consciousness is what is really behind the constructed interface we call reality, though it is formless and unobservable, undefinable by quantum physics' terms (Sarker 80). This "unknowable argument" resembles Gödel's incompleteness theorem (Gödel 38), which suggests that our knowledge will always be limited (Thuan and Ricard 35). In this case, we really have no way of objectively looking at the universe, as we are part of the universe. We will also bring to mind the Heisenberg Uncertainty Principle, which literally tells us, much like Gödel's theorem, that we cannot know all of a system all at one time (Heisenberg 32).

We have not yet answered our original question, whether samsara or vacuum genesis is more appropriate in defining our universe. However, various theorems and principles have all spoken of our understanding, or lack thereof, in describing true reality. Gödel's incompleteness theorem, Heisenberg's uncertainty principle, Donald Hoffman's Interface Theory of Perception, and the Advaita Vedanta all share the same underlying implication, all having revealed that our understanding of the universe is severely lacking. However, whatever that something is, we may never know it by looking through the lenses we use now.

Conclusion

Buddhism and quantum physics each carry implications for the meaning and mystery of true reality. I have argued that many important intersections exist between the two. I showed that, the Heisenberg Cut and the Buddhist concept of vikalpa, both suggest that mental boundaries are not key to reality, just constructs of the mind. Turning to Yogācāra and Donald Hoffman's Interface Theory of Perception, I argued both challenge the conventional view of what we see is, instead, that it is merely a construct. In this way, vikalpa becomes a veil for obscuring this deeper, fundamental true reality. Similarly, Pratītyasamutpāda and David Bohm's Implicate Order continue this line of thinking, rejecting the view that the universe consists of isolated parts. Instead, pratītyasamutpāda and David Bohm's Implicate Order showed a reality full of dynamism and impermanence, that true reality is deeply connected, behind the mental constructs of vikalpa. Finally, the concepts of saṃsāra and vacuum genesis remind us that our understanding of the cosmos may always be incomplete—that the full scope of reality might remain, in the end, ineffable.

Works Cited

Bell, Zebedee.

"The Holographic Principle." *Young Scientists Journal*, https://www.youngscientistsjournal.com/post/the-holographic-principle

Bodhi, Bhikkhu.

In the Buddha's Words: An Anthology of Discourses from the Pāli Canon. Urban Dharma, https://urbandharma.org/pdf2/InTheBuddhasWordsBookPreview.pdf.

Bohm, David.

Wholeness and the Implicate Order. Routledge, 1980. https://www.taylorfrancis.com/books/mono/10.4324/9780203995150/wholeness-implicate-order-david-bohm.

Borzeszkowski, Horst-Heino., and Renate Wahsner

"Heisenberg's Cut and the Limitations of Its Movability" *Annalen der Physik*, vol. 45, no. 48, 1988, pp. 522–528, https://dpl6hyzg28thp.cloudfront.net/media/Annalen_der_Physik_-_1988_-_v._Borzeszkowski_-_Heisenberg_s_Cut_and_the_Limitations_of_i_v3PYPEd.pdf

Chu, Elise L.

"A Dialogue Between Buddhism and Quantum Physics." *ResearchGate*, 2019, https://www.researchgate.net/publication/335576699_A_Dialogue_Between_Buddhism_and_Quantum_Physics.

Gödel, Kurt.

"On Formally Undecidable Propositions of Principia Mathematica and Related Systems." Translated by B. Meltzer, 1931. *Homepages.uc.edu,* https://homepages.uc.edu/~martinj/History_of_Logic/Godel/Godel%20–%20On%20Formally%20Undecidable%20Propositions%20of%20Principia%20Mathematica%201931.pdf

Griffiths, David J.

Introduction to Quantum Mechanics. 2nd ed., Pearson Prentice Hall, 2005. https://www.fisica.net/mecanica-quantica/Griffiths%20-%20Introduction%20to%20quantum%20mechanics.pdf

Heisenberg, Werner.

"The Actual Content of Quantum Theoretical Kinematics and Mechanics." Translated by Arnold J. Kálnay, NASA Technical Memorandum 77379, NASA, 1983. Originally published 1927. https://ntrs.nasa.gov/citations/19840008978.

Hincks, Adam D.

"What Does Physical Cosmology Say about Creation from Nothing?" *Creation ex Nihilo*, edited by Gary A. Anderson and Markus Bockmuehl, University of Notre Dame Press, 2018, pp. 319–346.

Hoffman, Donald D., and Chetan Prakash.
"Objects of Consciousness." *Frontiers in Psychology*, vol. 5, 2014, https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00577/full

Hoffman, Donald D., et al.
"The Interface Theory of Perception." *Psychonomic Bulletin & Review*, vol. 22, 2015, pp. 1480–1506, https://link.springer.com/article/10.3758/s13423-015-0890-8.

Jaeger, Lars.
The Eternal Interplay: Surprising Overlaps Between Quantum Physics and Buddhism. Digilib Stekom, 2022, https://digilib.stekom.ac.id/assets/dokumen/ebook/feb_1d54b7880a60c738090c798ba2c3472126c4e44a_1646976695.pdf.

Ling, Samuel J., Jeff Sanny, and William Moebs.
University Physics Volume 3. OpenStax, 2016. https://openstax.org/books/university-physics-volume-3/pages/1-introduction

Narasimhan, Ashok., et al.
"The Nature of the Heisenberg-von Neumann Cut: Enhanced Orthodox Interpretation of Quantum Mechanics." *Activitas Nervosa Superior*, vol. 61, 2019, pp. 12-17 https://link.springer.com/article/10.1007/s41470-019-00048-x.

Ricard, Matthieu, and Trinh Xuan Thuan.
The Quantum and the Lotus: A Journey to the Frontiers Where Science and Buddhism Meet. Three Rivers Press, 2004. https://archive.org/details/thequantumandthelotuswherescienceandbuddhismmeetmricardtxuanthuan_764_m/page/n31/mode/2up.

Sarker, Chitta Ranjan.
"Consciousness in Quantum Physics and Meaning in the Advaita Philosophy of Adi Sankaracharya." *Ultimate Reality and Meaning*, vol. 38, no. 1–2, 2015, pp. 73–81. University of Toronto Press, https://utppublishing.com/doi/full/10.3138/uram.38.1-2.73.

Sopa, Geshe Lhundub.
Steps on the Path to Enlightenment: A Commentary on Tsongkhapa's Lamrim Chenmo, vol. 2: *Karma*, edited by David Patt, Wisdom Publications, 2005. https://books.google.com/books?id=LkWMzwEACAAJ&newbks=0&hl=en&source=newbks_fb.

Tabaczek, Mariusz.
"Vacuum Genesis and Spontaneous Emergence of the Universe from Nothing in Reference to the Classical Notion of Causality and Creation ex nihilo." *ResearchGate*, 2024, https://www.researchgate.net/publication/381490205.

Tegchok, Khensur Jampa.
Insight into Emptiness. Wisdom Experience, 2012, pp. 17-31. https://wisdomexperience.org/academy/wp-content/uploads/sites/4/2017/02/Buddhism-Lesson-5-Reading-2.pdf.

Tischler, Ethan.

Emptiness and Wholeness: Untangling the 'Realities' of Tibetan Buddhism and Quantum Physics. 2014. Wesleyan University, CiteseerX, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=382ac47a3091eecf03f203378e64a17873180395.

Waldron, James. *The Co-Arising of Mind and World*. Middlebury College, 2006, pp. 175-208. https://www.middlebury.edu/college/sites/default/files/2023-03/waldron_co-arising_of_mind_and_world0.pdf.

Wang, George S.

"On the Consistency of Modern Physics and Buddhism with Oneness as the Main Subject of Discussion." *Hsi Lai Journal of Humanistic Buddhism,* vol. 8, 2007, http://www.modb.url.tw/content/article/E8-On%20the%20Consistency%20of%20Mod[...]0subject%20of%20discussion.pdf

Postpartum Depression – A Mother's Burden or Society's Neglect?
By Jira Wongswan

Abstract

Postpartum depression is a mental illness that affects mothers all over the world, yet there is generally a lack of research into the factors of its root. Certain areas of the world recognise this issue and are using various approaches to mitigate it; whereas others produce little research on the topic. The research process has identified the social determinants of postpartum depression to be recognised in certain regions globally, whereas not in others. Greater recognition of the role social factors play in influencing the likelihood of symptom development could improve the acknowledgement and treatment of the disorder, and even support the introduction of preventive measures. Research into this has highlighted the need for further research in lower- to lower middle- income countries where less research was available on the topic of causation compared to other areas of the world.

Introduction

Postpartum depression is defined as a condition where a mother may experience emotional swings, frequent crying, fatigue, guilt and other symptoms commonly associated with depression after giving birth (Cleveland Clinic, 2022). Affecting many women globally, there are multiple factors that come into consideration on the overall length and severity of each patient's symptoms.

Social determinants of health refer to the environment people are born in and grow up, live, work and age, as well as the system created to deal with the illnesses (World Health Organization, 2013). The potential social determinants of postpartum depression will be discussed and how they may have changed over time. Jeste (2022) explains that The World Health Organisation (WHO) defines social determinants of health (SDoH) as the 'non-medical factors' that play a role in healthcare decisions. Furthermore, it is crucial to mention that this paper is based primarily on Western World research, especially in the UK and US, due to the lack of published literature on the topic elsewhere in the world.

There have been various sources regarding social determinants of health having some responsibility towards the development of conditions such as postpartum depression worldwide (Williams, 2003). Specifically, a 2023 paper in Vietnam states that despite the various sources pointing the determinants towards many mental and physical health problems, postpartum depression continues to be misdiagnosed. This is possibly due to a lack of recognition of the signs of the disorder (Nguyen et al., 2023). Interestingly though, it was hypothesized by a study conducted in Indonesia in 2024, that those who weren't around their family as much throughout the whole pregnancy would exhibit symptoms of postpartum depression (Mahmudah et al., 2024). Therefore, there seems to be a shared consensus amongst literature that there is a relationship between the presence of other people and the likely onset of postpartum depression (Agrawal et al., 2022; Corrigan et al., 2015). However, although these cited papers are from the

continent of Asia, there remains a need for more research focusing on non-Western areas of the world (Low et al., 2023; Wang et al., 2021). Despite conducting a deep search of literature using various search terms, it was not possible to gather a range of material from many areas of the world outside of those already mentioned.

In this paper, the term Postpartum Depression (PPD) will be used to refer to the period after a mother has given birth where she may experience emotional highs and lows, frequent crying, and other common symptoms of depression (Cleveland Clinic, 2022). The term is sometimes used interchangeably with "postnatal depression". The term "postpartum" garnered results from the US, while "postnatal" was the term used that provided literature from the UK.

Predisposing factors

Some women may be more susceptible to developing PPD due to factors that have exposed them to vulnerability to developing the condition, known as predisposing factors (Monje et al., 2023; Weingarten et al., 2021). For example, pre-existing mental disorders such as Bulimia Nervosa (BN) and Binge Eating Disorder (BED) can affect the possibility of developing PPD and how severe it may present (Mazzeo et al., 2006). This is because the presence of pre-existing vulnerabilities may cause PPD to develop, or encourage other diagnoses associated with Anorexia Nervosa (AN) to appear, such as an unhealthy reduction in food intake to maintain a certain body shape (Makino et al., 2020). A reason for this occurring can be due to hormonal imbalance, leading to feelings of criticism or shame. This is essential to consider amongst the mental health of the mothers, as this vulnerability may trigger the disorder to re-present itself. Moreover, the same can be said with those with a history of depression (Liu et al., 2022; Zhao & Zhang, 2020 as cited in Agrawal et al., 2022). Additionally, non-medical factors e.g. location in the world, can influence a doctor's diagnosis, which could potentially affect the treatment. This may provide insight as to why some patients with the same symptoms receive varying levels of support and response to their condition (McKinlay et al., 1996).

There are multiple factors that can affect PPD that can be within a mother's external everyday environment. An example is the mother's marital status; for example, if they are in a happy and stable marriage, this dynamic may help to soothe and lessen the symptoms of PPD. On the other hand, those mothers in an unhappy, or even abusive marriage, may find their PPD worsening due to the excess mental toll felt from the relationship difficulties (Agrawal et al., 2022). Moreover, women who grow up in supportive, loving families may find it easier to confide in others about their issues. It is worth considering that those raised in households where traditional ideologies prevail—such as the belief that women are inferior to men—may feel unable to express concerns about their mental wellbeing. This could cause further isolation for the woman and a worsening of symptoms. The same could be said for the presence of friends during the postpartum period, as increased social support is generally considered to be positively correlated with milder and less symptomatic PPD (Corrigan et al., 2015).

Research has found that the number of children the mother has can affect the PPD. This may be because mothers with more children may be more preoccupied with parenting, so may

not have time or resources to best care for their own mental health (Cho et al., 2022). A study in 2022, by Cho et al. defined social support as a 'beneficial exchange of materials among people'. They suggested that experienced mothers may receive less social support, due to others perceiving them as capable, leaving them to feel alone; potentially worsening postpartum symptoms. Inexperienced, first-time mothers may go through the same. They may feel unsure on how to proceed with the aftercare of having a baby and face difficulties with knowing how best to do things. Due to this overwhelming time, symptoms of low mood or anxiety may occur, which without mental health intervention, could potentially lead to an eventual diagnosis of PPD. Like raising a child, the environment the woman is surrounded by can affect her experience pre-, post-, and during birth. There is a common analogy that when a sunflower fails to grow, no one blames the sunflower solely for it, but rather the environment it's surrounded by. People will usually water the flower more or feed it with fertilizer to induce more growth, because it's not the sunflower's fault for its inability to grow. The same can sometimes be said with many mental health issues. Mothers in a warm and loving environment may still have a high chance of experiencing PPD, but a supportive system can help make recovery easier and quicker as well as reducing the possibility of getting PPD to begin with (Cho et al., 2022).

Global determinants of PPD

Depending on where the mother is located in the world, the societal pressure placed on her can be different. This could be regarded as factors that affect her risk of getting PPD like predisposing factors (Monje et al., 2023). For example, some cultures e.g. Turkey, may impose unrelenting high standards for the mother's roles in the house (Durak et al., 2023); Kaplan, 2023). This could be placing pressure on the mother to hold a job outside as well as conform to the role of a mother at home. While their job may already be emotionally and physically demanding, having to do so after having a child and when hormones are imbalanced and dysregulated may place further stress on the new mother (Masih & Masih, 2022).

Furthermore, the development of a country also affects a mother's overall experience with PPD. The more developed a country and higher economic growth there is, the more advanced their healthcare may be, due to a higher "proportion of healthcare expenditure in total GDP" also increasing (Bedir, 2016). While this helps to address physiological symptoms, it may even help the mother mentally as they may meet other women going through similar experiences to them. In the Western world, healthcare is more advanced so there has been more progress in overall natal care compared to the rest of the world (Yu & Dunn, 2004). This is because many countries in the EU and commonwealth countries use public funding to provide access to quality healthcare to the poorest (Thomas, 2023). This is essential in a mother's postpartum symptoms, as early detection can help to ensure the mother heals after the pregnancy mentally. Such healthcare systems can also help during the pregnancy to ensure the mother's mental health is cared for, which can even prevent PPD developing (Campbell & Graham, 2006 as cited in World Health Organization, 2016). This was discovered through research from Western areas, as other areas generally lacked literature regarding this topic.

How has time played a factor into the presence of PPD

There are several factors that influence PPD. Some may argue that we could have potentially caused PPD due to the modern social structure, while some may believe that PPD has always been present, only disregarded (Hahn-Holbrook & Haselton, 2014).

Lifestyle factors such as dietary intake could be relevant to factors that affect the presence of PPD. In medieval times, food wasn't as varied as the present (Vermeulen et al., 2020). Nowadays, food is more varied and accessible, with many options catering for those following a range of different dietary needs. There has also been an increase in the provisions of more unhealthy food, which are perhaps more accessible to some than those healthier options. These choices can worsen peoples' overall health, which we can think about in the context of PPD. It has been proven that there is a decline in nutritional quality in various foods, which require more attention redirected at them due to them being crucial to an individual's health (Bhardwaj et al., 2024; Fernandes et al., 2023).

Like many mental disorders, PPD is difficult to represent accurately as the experiences vary for every mother. Written in 1890, by Charlotte Perkins Gilman, "The Yellow Wallpaper" depicts a woman's experience with being placed on bed rest to cure her PPD. Although it was written a considerably long time ago, it manages to highlight how being caged in a room was possibly the wrong approach for the situation. In the book, the woman is placed on bed rest by her husband, who is a General Practitioner (GP). It feels important to note that this could perhaps reflect the lack of development in postnatal care in the 1800s and the roles of women and men in society at that time. This is highlighted throughout the book by the husband's overall disbelief in the woman's concerns, as well as the lack of compassion and empathy he shows her.

Another argument in support of society playing a considerable role in the cause of PPD, is social media. While it has made globalization and contacting your loved ones easier, it is worth considering the disadvantages. Social media has created multiple new insecurities for teenage girls especially (Berne et al., 2014 cited in Popat and Tarrant, 2023) and has created an environment where harassment and bullying thrive (O'Reilly et al., 2019; Singleton et al., 2016 as cited in Popat and Tarrant, 2023). Such experiences have been highlighted by the recent release of a UK TV show called *Adolescence*, during which the main character prides himself for not physically assaulting a classmate when she rejected him. This show has triggered much controversy, including questions as to whether social media is healthy for the growing youth (Hogan, 2025). Although there can be useful outcomes from the use of social media, the negative impacts may be more detrimental to developing youths. This is supported by the work of United States Surgeon General Vivek Murthy, MD, MBA, who suggests that the use of social media is "disrupting important healthy behaviors", e.g. scrolling on their phone instead of going outside and reading a book. The extent of this damage depends on the time spent on social media, but a study conducted by WHO in September 2024, found that more than 10% of adolescents found themselves "struggling to control their use and experiencing negative consequences" (Berne et al., 2014).

Nowadays, unfortunately, the media will often prioritise demonstrating the different unjust acts committed when a mother is suffering through PPD, as opposed to trying to clarify and destigmatise the difficult journey of PPD (Goc, 2004). Therefore, as a step forward, social media could provide more educational information regarding the whole pregnancy process. A possible approach for this could be users and companies to monitor the representation of mental disorders like PPD more regularly; this could be done by helping to boost posts that do spread awareness about these disorders properly, as well as reporting content that could be spreading misinformation.

This overuse of technology can negatively impact mothers greatly, a study by Zhang and colleagues in 2020 found that new mothers who take selfies, and share more of their personal life online, tend to develop PPD (Zhang et al., 2020 as cited in Ye, 2023). A 2021 study conducted in Germany found that social media played an integral role in "postpartum body dissatisfaction and eating disorder psychopathology" (Nagl et al., 2021). They measured the average of frequent use of apps like Facebook and Instagram, per week. This was used alongside the Body Shape Questionnaire to measure perception of body image. The Eating Disorder Examination Questionnaire was to detect potential issues with an individual's eating behaviors (Jennings & Phillips, 2017). By conducting these measurements, a general pattern appeared: during the postpartum period, those using social media more frequently experienced more body image dissatisfaction. This was theorized to be due to two main components: appearance-related comparisons and thin-ideal internalization (Nagl et al., 2021). Therefore, it could be concluded that social media use during pregnancy could be harmful towards mothers, as the entire pregnancy period can be an emotionally and mentally challenging time.

It is worth noting the positive effects of the changes in society. An example is social media, which can have a positive impact on mothers, as many found that it provided emotional support that helped them when going through a difficult time (Kumarasinghe et al., 2024). Since PPD can be a life-threatening condition, it is essential that the condition is dealt with appropriately and timely, for the health of the mother and baby (Amiri & Behnezhad, 2021). This is why it's essential that social factors like social media are recognized as potential impacts towards the mother's mental health; both positively and negatively. Due to the widespread popularity of the internet, people are able to share opinions and resources regarding PPD. This large pool of data can increase PPD detection and treatment (Ye, 2023). Therefore, a possible approach to integrate social media healthily into the postpartum period, is to provide a space where mothers can contact each other and discuss hardships together. The key being to encourage a supportive environment, and to celebrate one-another's successes. This can help create a sense of community, offering support throughout what can be a challenging period.

Acknowledgement of PPD within society

In some areas around the world, women are often told to grow up, bear children, and become a housewife (Mussida & Patimo, 2021). This creates a society where men tend to have jobs and earn income, while women take care of the house and children. While not always true,

this can cause some households to regard the man as the "man of the house", emphasizing the man's superiority over the woman (Anwar et al., 2022). Therefore, this can cause the man to joke or act condescending towards the woman's concerns – both mentally and physically. Additionally, another group that possibly are unfamiliar with the term PPD and its effects, are those with an immigrant background (Chien et al., 2012). This can be both men and women, due to the culture created that can cause mental disorders like depression to be perceived as madness (Dennis & Chung-Lee, 2006; Schmied et al., 2017 as cited in Johansson et al.,2023).

Additionally, the lack of knowledge regarding PPD and other disorders caused by pregnancy has led to an inconsistent length of maternity leaves. Women in the UK are subject to the regular 26 weeks, 182 days of maternity leave, before being allowed an additional 26 weeks (NHS UK, 2020; UK Government, n.d.). However, women in Thailand are subject to 98 days of maternity leave if in the private and public sector. Although, those in the public sector could be offered an additional 90 days of leave at half their wage (Amanda Chua, 2025). Therefore, this comparison could perhaps reveal a pattern of Lower Income Countries (LICs), usually allowing shorter maternity leaves due to lower overall Gross Domestic Product (GDP) and Gross National Income (GNI). This raises the issue of whether there is too much emphasis from companies on having low production costs, possibly causing them to disregard the welfare of their workers. Moreover this problem seems to be rather prevalent in all areas of the world, as even in a High-Income Country (HIC) like the US, there is no national maternity or paternity leave programme (Arneson, 2021). Therefore, this perhaps brings to light a global issue that may need to be resolved to reduce a risk factor for PPD.

Developments in perinatal care

During this research process various search terms were used to collect information on depression around the time of pregnancy and following. Due to different terms being used in different areas of the world to describe this period, it highlights the discrepancies in search results when researching these topics. While still a developing topic in the healthcare world, there is ongoing research into psychological care during the prenatal period, which some countries address using perinatal mental health services. In the UK, perinatal care generally refers to the mental assistance to mothers during the pregnancy or up to a year after the birth (NHS England, n.d.; Webb et al., 2021). Additionally, the terms antenatal and prenatal care essentially mean the same thing: the period before birth when a woman receives healthcare and support (National Health Service UK, 2020; World Health Organization, 2016). Postpartum and postnatal care refer to the medical care given to women up to 8 weeks after giving birth (Lopez-Gonzalez & Kopparapu, 2022). This was considered an especially pressing concern, as identified by Buultjens and colleagues in 2021, that poor mental health played a prominent role in the cause of morbidity in mothers globally. Therefore, it was found that any support socially helped to reduce the effect of depression (Buultjens et al., 2021).

In the UK, National Institute of Health and Care Excellence (NICE) guidelines state that women with "a higher risk of adverse outcomes" are subject to antenatal care by 10 weeks.

Ongoing relevance for these communities, first stated in 2012, updated in 2023, shows that women higher risk of adverse outcomes: Black, Asian (except Chinese) and mixed ethnic backgrounds living in the most deprived areas generally access antenatal care late (NICE, 2012). Some global locations, for example in the UK, place importance on providing antenatal care to those facing more adversity. Whilst some global locations, usually LICs e.g. Malawi, could have a larger population falling into this category, do not publish information on higher-risk women amongst healthcare prioritisation (Townsend et al., 2021; Wellcome Trust, 2023). This of course does not mean these prioritisations are not necessary, but may be more a sign of a lack of published research in these areas. There have been milestones in the progression of perinatal care, e.g. the "establishment of national training programmes" (Yu & Dunn, 2004), which are programmes established to help train medical professionals to identify high risk mothers and infants to ensure that the hospital is prepared. Many countries are still behind in the overall psychological aspect of perinatal care, as areas of higher development generally will have made advancements in perinatal care, like the USA, the UK, and Australia (Yu & Dunn, 2004).

The research into postnatal care is ever-growing, with some researchers highlighting postnatal care's importance on reducing the child and maternal mortality and morbidity rates (McCauley et al., 2022). This is due to the Lower Middle-Income Countries (LMIC) e.g. Ghana (OECD, 2023), generally having to encounter more obstacles to access healthcare, therefore perhaps forcing them to disregard maternal mental health. This may be due to the overall lack of access to quality of healthcare within the country (Dawkins et al., 2021). This is similar to the point made earlier regarding places such as Malawi. It was found during the research process that generally LMIC e.g. Kenya (OECD, 2023) to LICs e.g. Sierra Leone (Wellcome Trust, 2023), may lack the ability to provide sufficient perinatal care to mothers in higher risk situations (Mohseni et al., 2023). This is a pressing concern as it raises the question on whether it is fair that some mothers will have more access to quality healthcare compared to others. Especially with regards to the potential significance towards their mental and physical state after giving birth.

Highlighted through an article by McCauley and her colleagues in 2022, there are 14 components that could be used when providing postnatal care. This includes screening and management of the mental health of mothers, which could help to prevent or reduce PPD symptoms. Additionally, while general perinatal care can be seen as rather undeveloped compared to other branches of healthcare, there has been progression in the amount of attention redirected to women with "complex social factors" (NICE, 2010). Examples of possible "complex social factors" regarded by NICE in the UK, include homelessness, poverty, misuser of substances, asylum seeker, under 20 years of age, and many more. Mothers that fall into these categories will be provided with extra provision and care.

Furthermore, as well as more care provided to those with higher social risk factors, there has also been more guidance for women with pre-existing physical illnesses e.g. diabetes (NICE, 2016a). This may be due to the spike in blood glucose levels that occur during pregnancy, creating a need for more advisory for those with type 1 diabetes. This type of diabetes is caused

by the immune system preventing production of insulin, an important hormone that reduces the blood glucose level when it gets too high. The absence of insulin causes glucose to build up in the bloodstream which can cause a multitude of physiological damage if not treated properly (Mayo Clinic Staff, 2024). In the UK, pregnant women are offered numerous additional services, for example constant monitoring of their blood glucose levels and testing of blood glucose after birth. Additionally, women with gestational diabetes, who tested negative for diabetes after birth as their blood glucose levels dropped back to their normal range, are offered yearly HbA1c tests. NICE provides this treatment to determine the risk of the woman getting type 2 diabetes, known as late on-set diabetes, and the possibility of undiagnosed type 1 diabetes that went unnoticed (NICE, 2016b).

Overall, from conducting research in this area, there appears to be a wide healthcare disparity for mothers of different backgrounds. This is supported by a literature review by Dagher and Linares in 2022, where it was found that Black mothers generally had the worst healthcare outcomes globally. This meant they had higher mortality rates, which Dagher and Linares theorise is due to structural racism present in their daily lives. Therefore, there is a clear need for more research into antenatal care at each step of the journey, including understanding about a woman's social surroundings. The approach to developing antenatal care needs careful consideration towards the different experiences of every mother, in various locations across the world. Necessary information should be collated on the woman's life, besides medical-based aspects alone (Carroli et al., 2001). This paper's research highlights the importance of considering a woman's social environment, along with other aspects of her daily life, as these factors can significantly impact her mental health and influence the risk of experiencing PPD symptoms.

Conclusion

Overall, the research process found that there was a significant lack of research on the topic found in non-Western areas of the world, such as Africa and Asia. As mentioned, various search terms were used to avoid geographical bias in search results. This involved using a range of terminology referring to the same topics e.g. postpartum is widely used in the US, while postnatal is widely used in the UK. It was possible to find information on the perception, diagnosis and treatment for post-natal care in certain global areas, including how these may be changing with time. Improvements included more acknowledgement of the many contributing factors to the condition, such as a woman's external environment and support network around her.

This research process has revealed a significant lack of published studies on postnatal care across many regions of the world, highlighting the need for further research to improve public health outcomes.

Works Cited

A. Saha & M. Das. (2017). Impact of social networking sites on post-partum depression in women: An analysis in the context of Bangladesh. 2017 20th International Conference of Computer and Information Technology (ICCIT), 1–6. https://doi.org/10.1109/ICCITECHN.2017.8281831

Acharya, K. P., Ghimire, T. R., & Subramanya, S. H. (2021). Access to and equitable distribution of COVID-19 vaccine in low-income countries. Npj Vaccines, 6(1), 54. https://doi.org/10.1038/s41541-021-00323-6

Aderoba, A. K., & Adu-Bonsaffoh, K. (2022). Antenatal and Postnatal Care. Obstetrics and Gynecology Clinics, 49(4), 665–692. https://doi.org/10.1016/j.ogc.2022.07.005

Agrawal, I., Mehendale, A. M., & Malhotra, R. (2022). Risk Factors of Postpartum Depression. Cureus, 14(10), e30898. https://doi.org/10.7759/cureus.30898

Allen, J., Balfour, R., Bell, R., & Marmot, M. (2014). Social determinants of mental health. International Review of Psychiatry, 26(4), 392–407. https://doi.org/10.3109/09540261.2014.928270

Alves, C., Soares, L., & Faria, A. L. (2023). Pregnancy and postpartum depression (PPD): A non-systematic review assessing the current state of knowledge on PPD. Revista Portuguesa de Investigação Comportamental e Social: RPICS, 9(1), 5.

Amanda Chua. (2025, April 16). Maternity Leave in Thailand: What You Need to Know - Omni HR. https://www.omnihr.co/blog/maternity-leave-thailand

Amiri, S., & Behnezhad, S. (2021). The global prevalence of postpartum suicidal ideation, suicide attempts, and suicide mortality: A systematic review and meta-analysis. International Journal of Mental Health, 50(4), 311–336. https://doi.org/10.1080/00207411.2021.1959814

Amy Marcus-Newhall, Bettina Casad, Judith Lemaster, Jennifer Peraza, & Nicole Silverman. (2009). Mothers' Work‐Life Experiences: The Role of Cultural Factors. In ResearchGate. https://www.researchgate.net/publication/279370372_Mothers'_Work-Life_Experiences_The_Role_of_Cultural_Factors

Anwar, S. M., Sardar, Q. A., & Khan, A. A. (2022). A Sociological Analysis of Employment Discrimination Against Women in Pakistan. Journal of Management Practices, Humanities and Social Sciences, 6(1), Article 1. https://doi.org/10.33152/jmphss-6.1.2

Astrachan-Fletcher, E., Veldhuis, C., Lively, N., Fowler, C., & Marcks, B. (2008). The Reciprocal Effects of Eating Disorders and the Postpartum Period: A Review of the Literature and Recommendations for Clinical Care. Journal of Women's Health, 17(2), 227–239. https://doi.org/10.1089/jwh.2007.0550

Aved, B. M., Irwin, M. M., Cummings, L. S., & Findeisen, N. (1993). Barriers to prenatal care for low-income women. Western Journal of Medicine, 158(5), 493–498.

Bedir, S. (2016). Healthcare Expenditure and Economic Growth in Developing Countries. Advances in Economics and Business, 4(2), 76–86. https://doi.org/10.13189/aeb.2016.040202

Berne, S., Frisén, A., & Kling, J. (2014). Appearance-related cyberbullying: A qualitative investigation of characteristics, content, reasons, and effects. Body Image, 11(4), 527–533. https://doi.org/10.1016/j.bodyim.2014.08.006

Bhardwaj, R. L., Parashar, A., Parewa, H. P., & Vyas, L. (2024). An Alarming Decline in the Nutritional Quality of Foods: The Biggest Challenge for Future Generations' Health. Foods, 13(6), 877. https://doi.org/10.3390/foods13060877

Bloch, M., Rotenberg, N., Koren, D., & Klein, E. (2006). Risk factors for early postpartum depressive symptoms. General Hospital Psychiatry, 28(1), 3–8. https://doi.org/10.1016/j.genhosppsych.2005.08.006

Brown, J. D., Harris, S. K., Woods, E. R., Buman, M. P., & Cox, J. E. (2012). Longitudinal Study of Depressive Symptoms and Social Support in Adolescent Mothers. Maternal and Child Health Journal, 16(4), 894–901. https://doi.org/10.1007/s10995-011-0814-9

Buultjens, M., Farouque, A., Karimi, L., Whitby, L., Milgrom, J., & Erbas, B. (2021). The contribution of group prenatal care to maternal psychological health outcomes: A systematic review. Women and Birth, 34(6), e631–e642. https://doi.org/10.1016/j.wombi.2020.12.004

Campbell, O. M. R., & Graham, W. J. (2006). Strategies for reducing maternal mortality: Getting on with what works. Lancet (London, England), 368(9543), 1284–1299. https://doi.org/10.1016/S0140-6736(06)69381-1

Carroli, G., Rooney, C., & Villar, J. (2001). How effective is antenatal care in preventing maternal mortality and serious morbidity? An overview of the evidence. Paediatric and Perinatal Epidemiology, 15(s1), 1–42. https://doi.org/10.1046/j.1365-3016.2001.0150s1001.x

Chien, L.-Y., Tai, C.-J., & Yeh, M.-C. (2012). Domestic Decision-Making Power, Social Support, and Postpartum Depression Symptoms Among Immigrant and Native Women in Taiwan. Nursing Research, 61(2), 103. https://doi.org/10.1097/NNR.0b013e31824482b6

Cho, H., Lee, K., Choi, E., Cho, H. N., Park, B., Suh, M., Rhee, Y., & Choi, K. S. (2022). Association between social support and postpartum depression. Scientific Reports, 12(1), 3128. https://doi.org/10.1038/s41598-022-07248-7

Cleveland Clinic. (2022, April 12). Postpartum Depression: Causes, Symptoms & Treatment. Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/9312-postpartum-depression

Corrigan, C. P., Kwasky, A. N., & Groh, C. J. (2015). Social Support, Postpartum Depression, and Professional Assistance: A Survey of Mothers in the Midwestern United States. The Journal of Perinatal Education, 24(1), 48–60. https://doi.org/10.1891/1058-1243.24.1.48

Dagher, R. K., & Linares, D. E. (2022). A Critical Review on the Complex Interplay between Social Determinants of Health and Maternal and Infant Mortality. Children, 9(3), Article 3. https://doi.org/10.3390/children9030394

Dahl, E., & Malmberg‐Heimonen, I. (2010). Social inequality and health: The role of social capital. Sociology of Health & Illness, 32(7), 1102–1119. https://doi.org/10.1111/j.1467-9566.2010.01270.x

Dawkins, B., Renwick, C., Ensor, T., Shinkins, B., Jayne, D., & Meads, D. (2021). What factors affect patients' ability to access healthcare? An overview of systematic reviews. Tropical Medicine & International Health, 26(10), 1177–1188. https://doi.org/10.1111/tmi.13651

Dennis, C.-L., & Chung-Lee, L. (2006). Postpartum Depression Help-Seeking Barriers and Maternal Treatment Preferences: A Qualitative Systematic Review. Birth, 33(4), 323–331. https://doi.org/10.1111/j.1523-536X.2006.00130.x

Dyer, C. (1988). Changes in Diet in the Late Middle Ages: The Case of Harvest Workers. Bloomsbury Publishing Plc. https://doi.org/10.5040/9781472599407

Edwards, R. C., Thullen, M. J., Isarowong, N., Shiu, C.-S., Henson, L., & Hans, S. L. (2012). Supportive relationships and the trajectory of depressive symptoms among young, African American mothers. Journal of Family Psychology, 26(4), 585–594. https://doi.org/10.1037/a0029053

Foege, W. H. (2010). Social Determinants of Health and Health-Care Solutions. Public Health Reports®, 125(4_suppl), 8–10. https://doi.org/10.1177/00333549101250S403

Fuhrmann, D., Knoll, L. J., & Blakemore, S.-J. (2015). Adolescence as a Sensitive Period of Brain Development. Trends in Cognitive Sciences, 19(10), 558–566. https://doi.org/10.1016/j.tics.2015.07.008

Ghaedrahmati, M., Kazemi, A., Kheirabadi, G., Ebrahimi, A., & Bahrami, M. (2017). Postpartum depression risk factors: A narrative review. Journal of Education and Health Promotion, 6(1), 60. https://doi.org/10.4103/jehp.jehp_9_16

Goc, N. (2004). Mothers and Madness: The Media Representation of Postpartum Psychosis. Brill. https://doi.org/10.1163/9789004495371_008

Guevara, J. P., Morales, K., Mandell, D., Mogul, M., Charidah, T., Luethke, M., Min, J., Clark, R., Betancourt, L., & Boyd, R. (2023). Social Media-based Parenting Program for Women With Postpartum Depressive Symptoms: An RCT. Pediatrics, 151(3), e2022058719. https://doi.org/10.1542/peds.2022-058719

Hahn-Holbrook, J., & Haselton, M. (2014). Is Postpartum Depression a Disease of Modern Civilization? Current Directions in Psychological Science, 23(6), 395–400. https://doi.org/10.1177/0963721414547736

Hogan, M. (2025, March 17). 'Unnervingly on-the-nose': Why Adolescence is such powerful TV that it could save lives. The Guardian. https://www.theguardian.com/tv-and-radio/2025/mar/17/adolescence-netflix-powerful-tv-could-save-lives

Institute of Medicine (US) Committee to Study Outreach for Prenatal Care, & Brown, S. S. (1988). Barriers to the Use of Prenatal Care. In Prenatal Care: Reaching Mothers, Reaching Infants. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK217704/

Jennings, K. M., & Phillips, K. E. (2017). Eating Disorder Examination–Questionnaire (EDE–Q): Norms for Clinical Sample of Female Adolescents with Anorexia Nervosa,. Archives of Psychiatric Nursing, 31(6), 578–581. https://doi.org/10.1016/j.apnu.2017.08.002

Jeste, D. V. (2022). Non-medical social determinants of health in older adults. International Psychogeriatrics, 34(9), 755–756. https://doi.org/10.1017/S1041610222000709

Kariotis, T. C., Prictor, M., Chang, S., & Gray, K. (2022). Impact of Electronic Health Records on Information Practices in Mental Health Contexts: Scoping Review. Journal of Medical Internet Research, 24(5), e30405. https://doi.org/10.2196/30405

Kathy Katella. (2024, April 17). How Social Media Affects Your Teen's Mental Health: A Parent's Guide. Yale Medicine. https://www.yalemedicine.org/news/social-media-teen-mental-health-a-parents-guide

Kumarasinghe, M., Herath, M. P., Hills, A. P., & Ahuja, K. D. K. (2024). Postpartum versus postnatal period: Do the name and duration matter? PLOS ONE, 19(4), e0300118. https://doi.org/10.1371/journal.pone.0300118

Kuss, D. J., & Griffiths, M. D. (2011). Online Social Networking and Addiction—A Review of the Psychological Literature. International Journal of Environmental Research and Public Health, 8(9), Article 9. https://doi.org/10.3390/ijerph8093528

LaVeist, T. A., & Pierre, G. (2014). Integrating the 3Ds—Social Determinants, Health Disparities, and Health-Care Workforce Diversity. Public Health Reports (1974-), 129, 9–14. JSTOR.

Liu, X., Wang, S., & Wang, G. (2022). Prevalence and Risk Factors of Postpartum Depression in Women: A Systematic Review and Meta-analysis. Journal of Clinical Nursing, 31(19–20), 2665–2677. https://doi.org/10.1111/jocn.16121

Lopez-Gonzalez, D. M., & Kopparapu, A. K. (2022). Postpartum Care of the New Mother. In StatPearls. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK565875/

Low, S. R., Bono, S. A., & Azmi, Z. (2023). The effect of emotional support on postpartum depression among postpartum mothers in Asia: A systematic review. Asia-Pacific Psychiatry, 15(2–3), e12528. https://doi.org/10.1111/appy.12528

Mahmudah, M., Puspitasari, N., Sulistyorini, Y., & Marta, B. S. (2024). The Role of Family Social Support on Prenatal and Postpartum Depression in Surabaya City, Indonesia: A Multivariate Approach. Journal of Medicinal and Chemical Sciences, 7(9), 1118–1132. https://doi.org/10.26655/JMCHEMSCI.2024.9.2

Makino, M., Yasushi, M., & Tsutsui, S. (2020). The risk of eating disorder relapse during pregnancy and after delivery and postpartum depression among women recovered from eating disorders. BMC Pregnancy and Childbirth, 20(1), 323. https://doi.org/10.1186/s12884-020-03006-7

Marmot, M., Allen, J., Bell, R., Bloomer, E., & Goldblatt, P. (2012). WHO European review of social determinants of health and the health divide. The Lancet, 380(9846), 1011–1029. https://doi.org/10.1016/S0140-6736(12)61228-8

Mayo Clinic Staff. (2024, March 27). Type 1 diabetes—Symptoms and causes. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/type-1-diabetes/symptoms-causes/syc-20353011

Mazzeo, S. E., Slof‐Op't Landt, M. C. T., Jones, I., Mitchell, K., Kendler, K. S., Neale, M. C., Aggen, S. H., & Bulik, C. M. (2006). Associations among postpartum depression, eating disorders, and perfectionism in a population‐based sample of adult women. International Journal of Eating Disorders, 39(3), 202–211. https://doi.org/10.1002/eat.20243

McCauley, H., Lowe, K., Furtado, N., Mangiaterra, V., & van den Broek, N. (2022). Essential components of postnatal care – a systematic literature review and development of signal functions to guide monitoring and evaluation. BMC Pregnancy and Childbirth, 22(1), 448. https://doi.org/10.1186/s12884-022-04752-6

McDonald, L. M. (2018). Postpartum Depression in the News Media.

McKinlay, J. B., Potter, D. A., & Feldman, H. A. (1996). Non-medical influences on medical decision-making. Social Science & Medicine, 42(5), 769–776. https://doi.org/10.1016/0277-9536(95)00342-8

Mohseni, M., Isfahani, H. M., Moosavi, A., Mohammadian, E. D., Mirmohammadi, F., Ghazanfari, F., & Ahmadi, S. (2023). Health system-related barriers to prenatal care management in low- and middle-income countries: A systematic review of the qualitative literature. Primary Health Care Research & Development, 24, e15. https://doi.org/10.1017/S1463423622000706

Monje, A., Kan, J. Y., & Borgnakke, W. (2023). Impact of local predisposing/precipitating factors and systemic drivers on peri-implant diseases. Clinical Implant Dentistry and Related Research, 25(4), 640–660. https://doi.org/10.1111/cid.13155

Mousa, O., Sabati, S. Y. A., & Khars, F. A. (2023). Postpartum Depression: Neglected Issue in Maternal Health. International Journal of Nursing Education, 15(1), 72–77. https://doi.org/10.37506/ijone.v15i1.18995

Mussida, C., & Patimo, R. (2021). Women's Family Care Responsibilities, Employment and Health: A Tale of Two Countries. Journal of Family and Economic Issues, 42(3), 489–507. https://doi.org/10.1007/s10834-020-09742-4

Nagl, M., Jepsen, L., Linde, K., & Kersting, A. (2021). Social media use and postpartum body image dissatisfaction: The role of appearance-related social comparisons and thin-ideal internalization. Midwifery, 100, 103038. https://doi.org/10.1016/j.midw.2021.103038

National Health Service UK. (2020, December 2). Your antenatal care. Nhs.Uk. https://www.nhs.uk/pregnancy/your-pregnancy-care/your-antenatal-care/

Navarro, V. (2009). What we mean by social determinants of health. Global Health Promotion, 16(1), 05–16. https://doi.org/10.1177/1757975908100746

Nguyen, H. T. H., Hoang, P. A., Do, T. K. L., Taylor-Robinson, A. W., & Nguyen, T. T. H. (2023). Postpartum depression in Vietnam: A scoping review of symptoms, consequences, and management. BMC Women's Health, 23(1), 391. https://doi.org/10.1186/s12905-023-02519-5

NHS England. (n.d.). Perinatal mental health. Retrieved July 6, 2025, from https://www.england.nhs.uk/mental-health/perinatal/

NHS England. (2018, July 10). The past, present and future of innovation in mental health. NHS England Digital. https://digital.nhs.uk/blog/transformation-blog/2018/the-past-present-and-future-of-innovation-in-mental-health

NHS UK. (2020, December 1). Maternity and paternity benefits and leave. Nhs.Uk. https://www.nhs.uk/pregnancy/finding-out/maternity-and-paternity-benefits-and-leave/

NICE. (2010, September 22). Key priorities for implementation | Pregnancy and complex social factors: A model for service provision for pregnant women with complex social factors | Guidance | NICE. NICE. https://www.nice.org.uk/guidance/cg110/chapter/Key-priorities-for-implementation

NICE. (2012a, September 17). Quality statement 1: Access to antenatal care | Antenatal care | Quality standards | NICE. NICE. https://www.nice.org.uk/guidance/qs22/chapter/Quality-statement-1-Access-to-antenatal-care

NICE. (2012b, September 17). Quality statement 2: Risk assessment | Antenatal care | Quality standards | NICE. NICE. https://www.nice.org.uk/guidance/qs22/chapter/Quality-statement-2-Risk-assessment

NICE. (2016a, January 19). Overview | Diabetes in pregnancy | Quality standards | NICE. NICE. https://www.nice.org.uk/guidance/qs109

NICE. (2016b, January 19). Quality statement 5: Annual HbA1c tests | Diabetes in pregnancy | Quality standards | NICE. NICE. https://www.nice.org.uk/guidance/qs109/chapter/Quality-statement-5-Annual-HbA1c-tests

NICE. (2021, August 19). Overview | Antenatal care | Guidance | NICE. NICE. https://www.nice.org.uk/guidance/ng201

OECD. (2023). ODA recipients: Countries, territories, and international organisations. OECD. https://www.oecd.org/en/topics/oda-eligibility-and-conditions/dac-list-of-oda-recipients.html

O'Hara, M. W. (2013). Postpartum Depression: Causes and Consequences. Springer Berlin Heidelberg. https://books.google.co.th/books?id=2mm1BgAAQBAJ

O'Reilly, M., Dogra, N., Hughes, J., Reilly, P., George, R., & Whiteman, N. (2018). Potential of social media in promoting mental health in adolescents. Health Promotion International, 34(5), 981–991. https://doi.org/10.1093/heapro/day056

Popat, A., & Tarrant, C. (2023). Exploring adolescents' perspectives on social media and mental health and well-being – A qualitative literature review. Clinical Child Psychology and Psychiatry, 28(1), 323–337. https://doi.org/10.1177/13591045221092884

Postpartum Support International. (2024). What the Media Should Know: Myths & Mis-characterizations About Postpartum Depression and Related Illnesses.

Schmied, V., Black, E., Naidoo, N., Dahlen, H. G., & Liamputtong, P. (2017). Migrant women's experiences, meanings and ways of dealing with postnatal depression: A meta-ethnographic study. PLOS ONE, 12(3), e0172385. https://doi.org/10.1371/journal.pone.0172385

Shapiro, G. D., & Fraser, W. D. (2013). Review: Psychosocial and psychological interventions reduce postpartum depression. Annals of Internal Medicine, 159(4), JC8. https://doi.org/10.7326/0003-4819-159-4-201308200-02008

Silverman, M. E., Reichenberg, A., Savitz, D. A., Cnattingius, S., Lichtenstein, P., Hultman, C. M., Larsson, H., & Sandin, S. (2017). The risk factors for postpartum depression: A population-based study. Depression and Anxiety, 34(2), 178–187. https://doi.org/10.1002/da.22597

Singleton, A., Abeles, P., & Smith, I. C. (2016). Online social networking and psychological experiences: The perceptions of young people with mental health difficulties. Computers in Human Behavior, 61, 394–403. https://doi.org/10.1016/j.chb.2016.03.011

Stringfellow, C. M. (2016). How Does Social Media Impact the Postpartum Depression Experience.

Sunkel, C., & Sartor, C. (2022). Perspectives: Involving persons with lived experience of mental health conditions in service delivery, development and leadership. BJPsych Bulletin, 46(3), 160–164. https://doi.org/10.1192/bjb.2021.51

Thomas, L. (2023, February 25). Healthcare Systems Around the World. News-Medical. https://www.news-medical.net/health/Healthcare-Systems-Around-the-World.aspx

Townsend, R., Chmielewska, B., Barratt, I., Kalafat, E., Meulen, J. van der, Gurol-Urganci, I., O'Brien, P., Morris, E., Draycott, T., Thangaratinam, S., Doare, K. L., Ladhani, S., Dadelszen, P. von, Magee, L. A., & Khalil, A. (2021). Global changes in maternity care provision during the COVID-19 pandemic: A systematic review and meta-analysis. eClinicalMedicine, 37. https://doi.org/10.1016/j.eclinm.2021.100947

Vermeulen, S. J., Park, T., Khoury, C. K., & Béné, C. (2020). Changing diets and the transformation of the global food system. Annals of the New York Academy of Sciences, 1478(1), 3–17. https://doi.org/10.1111/nyas.14446

Wang, Z., Liu, J., Shuai, H., Cai, Z., Fu, X., Liu, Y., Xiao, X., Zhang, W., Krabbendam, E., Liu, S., Liu, Z., Li, Z., & Yang, B. X. (2021). Mapping global prevalence of depression among postpartum women. Translational Psychiatry, 11(1), 543. https://doi.org/10.1038/s41398-021-01663-6

Webb, R., Uddin, N., Ford, E., Easter, A., Shakespeare, J., Roberts, N., Alderdice, F., Coates, R., Hogg, S., Cheyne, H., Ayers, S., Clark, E., Frame, E., Gilbody, S., Hann, A., McMullen, S., Rosan, C., Salmon, D., Sinesi, A., … Williams, L. R. (2021). Barriers and facilitators to implementing perinatal mental health care in health and social care settings: A systematic review. The Lancet Psychiatry, 8(6), 521–534. https://doi.org/10.1016/S2215-0366(20)30467-3

Weingarten, S., Diop, S., Specht, C., Turmes, L., Juckel, G., & Mavrogiorgou, P. (2021). Differences in interactional behaviour in postpartum depression with and without pre-existing mental disorder. Comprehensive Psychiatry, 108, 152248. https://doi.org/10.1016/j.comppsych.2021.152248

Wellcome Trust. (2023). Low- and Middle-income Countries—List. Wellcome. https://wellcome.org/research-funding/guidance/prepare-to-apply/low-and-middle-income-countries

Williams, G. H. (2003). The determinants of health: Structure, context and agency. Sociology of Health & Illness, 25(3), 131–154. https://doi.org/10.1111/1467-9566.00344

World Bank. (n.d.). World Bank Country and Lending Groups. Retrieved July 6, 2025, from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

World Health Organisation. (n.d.). Social determinants of health. Retrieved July 25, 2025, from https://www.who.int/health-topics/social-determinants-of-health

World Health Organization. (2013, July 5). Social determinants of health: Key concepts. https://www.who.int/news-room/questions-and-answers/item/social-determinants-of-health-key-concepts

World Health Organization. (2016). WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience. In WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience. World Health Organization. https://www.ncbi.nlm.nih.gov/books/NBK409110/

World Health Organization. (2024, September 25). Teens, screens and mental health. https://www.who.int/europe/news/item/25-09-2024-teens--screens-and-mental-health

Xie, R.-H., He, G., Koszycki, D., Walker, M., & Wen, S. W. (2009). Fetal Sex, Social Support, and Postpartum Depression. The Canadian Journal of Psychiatry, 54(11), 750–756. https://doi.org/10.1177/070674370905401105

Ye, Z. (2023). Exploring the Effect of Social Media on Post-Partum Depression. Lecture Notes in Education Psychology and Public Media, 29(1), 75–80. https://doi.org/10.54254/2753-7048/29/20231401

Yu, V. Y. H., & Dunn, P. M. (2004). Development of regionalized perinatal care. Seminars in Neonatology, 9(2), 89–97. https://doi.org/10.1016/j.siny.2003.08.011

Zhang, W., Liu, L., Cheng, Q., Chen, Y., Xu, D., & Gong, W. (2020). The Relationship Between Images Posted by New Mothers on WeChat Moments and Postpartum Depression: Cohort Study. Journal of Medical Internet Research, 22(11), e23575. https://doi.org/10.2196/23575

Zhao, X., & Zhang, Z. (2020). Risk factors for postpartum depression: An evidence-based systematic review of systematic reviews and meta-analyses. Asian Journal of Psychiatry, 53, 102353. https://doi.org/10.1016/j.ajp.2020.102353

The Digital Touch : Exploring the Impact of AI's Role in Luxury Marketing on Consumer Purchase Intention By Yaana Karwa

Abstract

The integration of Artificial Intelligence holds promise and value in the luxury fashion industry. Although previous studies have extensively explored AI's implications in product design and creation, they predominantly focus on markets like the United States. Given the significant influence marketing has in determining consumer purchases it is essential to investigate this in greater detail while focusing on the rapidly growing Indian luxury market. Addressing the Indian market is important due to the lack of focus on India in existing studies. This paper explores how the integration of AI in generating marketing content for luxury fashion products affects consumer purchase intentions in India. It argues that AI-generated marketing content can noticeably influence consumer purchasing. The research methodology employs a qualitative approach, including a comprehensive literature review, content analysis of social media posts and brand websites, and interviews. The findings from this study will contribute to the debates on the effectiveness of AI-generated marketing strategies and campaigns influencing purchasing behavior. Moreover, this research will provide luxury fashion and marketing companies with future directions for optimizing their success in the field.

Keywords

Artificial Intelligence, luxury fashion, purchase intention, AI-generated marketing, visual marketing, consumer purchase behavior

Introduction

As we step into the "Digital Age," the impact of developing technology exemplified by Artificial Intelligence (AI) on the field of luxury fashion can be seen profoundly. Over the years, the luxury industry has been influenced by dynamic trends and ever-changing consumer behaviors. Hence, with the rise of Artificial Intelligence, the marketing sector in luxury fashion is also adapting to the trend. Previous studies have included detailed analyses on AI's impact in product design and use in luxury marketing; as a result, its influence on consumers' purchase intention remains unclear.

In this study, the research question that will be further explored is: *What is the impact of AI-generated marketing content and marketing strategy on consumer purchase intention towards luxury fashion apparel in India?* The Indian luxury market is projected to grow at a compound annual rate of 4.3% (Webology, 2023). Furthermore, studies have shown Indian consumers displaying a strong inclination toward luxury purchases (Mishra et al., 2021). Thereby, making our topic important to explore in the Indian subcontinent.

The research methodology will begin with a comprehensive literature review, providing valuable insights into relevant themes, supported by facts, data and statistics. Contributing to several insights on relevant themes and topics being obtained, making use of their facts, data and

statistics. Another important source of data will be the interviews conducted with accustomed luxury fashion consumers. Participants will be asked to respond to real-life examples of AI visual marketing content from luxury fashion brands, along with a series of questions, through which insights can be gained into their behavior as Indian consumers. Lastly, analyzing content via Instagram posts and marketing visuals of luxury brands will provide the study with an in-depth understanding of visual marketing strategy.

Knowledge of how AI integration in visual marketing affects consumer purchase intentions can be valuable for consumers, marketing agencies, and luxury fashion companies. Luxury purchases are thought to be significant decisions based on their perceived emotional and functional value (Xu and Mehta, 2022). Marketing is the key aspect that holds the power to make or break these decisions. Therefore, understanding the influence of AI on it can be helpful to improve company sales, influence, and growth worldwide. Acknowledging that the field of AI will only experience growth in the years to come, this paper provides an opportunity to make future plans and marketing objectives. This study further contributes to future researchers looking for effective marketing methods.

This paper is structured as follows: 1) Literature review, 2) Research Methodology, 3) Content analysis, 4) Interviews, 5) Discussion, 6) Implications, 7) Limitations and 8) Conclusion.

Research Methodology

To acquire detailed and insightful data, this study will employ exploratory qualitative methods. This research methodology utilizes a qualitative approach, encompassing an extensive literature review, content analysis of social media posts and brand websites, and interviews. To determine the selection of participants for data collection, purposive sampling was used by including candidates who have experience purchasing three or more luxury fashion products independently, and are well-versed in technology. Our target population aims to be luxury product consumers between the ages of 25-50 who can make financially independent purchase decisions.

Data provided in this research paper has all been collected via primary and secondary research methods. Interviews were the source for primary data, wherein the use of interview guides ensured an extensive topic list and overview was provided and interviewer bias was prevented. Secondary data collection was primarily done using the internet, through social media platforms like Instagram, website news articles, and previous research papers. The data analysis methods employed included thematic and content analysis, following the approach outlined by (Du et al., 2023). Data validity was verified using triangulation. With the use of multiple datasets and theories to address our research question, data accuracy and validity are assured.

Ethical practices such as voluntary participation, informed consent and anonymity of personally identifiable data were taken into account (Bhandari, 2021). Thereby, assuring unbiased results and participants' privacy involved asking open-ended questions that were neither biased or leading. Furthermore, questions were checked to ensure they were not

emotionally offensive towards participants. Protecting the anonymity of participants was of utmost importance, with their identities, names, and personal information remaining private.

To conclude, this methodology employs qualitative data collection methods through primary and secondary means. Data thematic and content analysis will be verified by triangulation. Lastly, measures have been taken to ensure this study is conducted ethically.

Literature Review

Previous studies have discussed multiple themes. This literature review focuses on a thematic approach, organizing findings around key themes such as AI-powered designing, brand reputation, marketed emotional values, and more. As AI becomes increasingly relevant and a common solution to problems across the world, understanding its implications in marketing practices is vital. Additionally, this study may provide actionable insights for luxury fashion businesses to enhance their marketing strategies.

First, the mention of consumer responses towards AI-designed or AI fashion products was a recurring theme across many studies. Liang et al (2019) demonstrated that consumers with higher involvement in fashion are more susceptible to using new technology. However, like general consumers, they remain unconvinced about buying a fashion AI product. Brüns and Meißner (2024) highlighted the negative reactions among followers and consumers of brands if they adopt GenAI for creating their social media marketing. Lee and Kim (2023) found that AI-designed clothing was viewed negatively compared to human-designed clothing due to differences between consumer and AI-generated fashion schemas. Overall, indicating a major negative response towards AI integration in fashion.

Second, positive responses and purchase intentions toward products when AI is part of the process in making or marketing a product was a key theme. Examples include AI customization, personalization, and creative assistance. In contrast to the first theme, Brüns and Meißner (2024) suggested that when GenAI aids humans with marketing content creation instead of replacing them entirely, negative reactions are reduced. Positive attitudes were noted in Liang et al (2019) based on the assistance provided by the AI product in helping consumers choose trendy outfits. Rahman et al (2023) advised luxury brands to use AI-Powered Digital Assistance (AIPDA) as a marketing tool on their online platforms due to its personalization benefits. Study 3 conducted by Kim and Lee (2023) further supported this theme by highlighting the positive attitudes towards AI customization in clothing, which improved consumer attitudes and purchase intentions.

Next, brand reputation concerning product authenticity, essence, and quality holds utmost importance. The marketing of a brand may be excellent; however, the brand's reputation itself plays a mediating role in affecting consumer response towards AI integration as well. Xu and Mehta (2022) showcase contrasting arguments for this theme. Negative reactions and diminished perceived brand essence result from using AI for designing luxury products with high emotional value. On the other hand, a functional brand essence highlighted through marketing can solve the

problem. Kim and Lee (2023) showed that human-designed clothing is favored over AI-designed due to higher perceived authenticity and quality in Studies 1 and 2.

Last, another common theme was the marketed emotional value of a product affecting purchase intention in relation to AI. Abdul et al (2021) studied the brand Manolo Blahnik and concluded that their followers had positive responses to social media content on Facebook fueled by optimistic emotions and feelings toward the brand. Xu and Mehta (2022) suggested AI technology is helpful in fashion by understanding human emotions to enhance customer interaction and create body-inclusive designs that generate positive feelings.

Table 1 : Literature Review Findings

Study	Theme	Methods	Country/ Market	Factors/ Dimensions/ Variables	Findings
1. Liang et al., 2019	Consumer response towards implementing an Artificial Intelligence product in the fashion industry	Quantitative	Developed countries (USA)	1. Consumers (gender and age group) 2. Perceived ease of use of AI device 3. Performance risk of AI device 4. Perceived usefulness of the AI device	1. Positive attitudes towards technology were seen to have a direct impact on influencing positive purchase intentions of consumers. 2. The study proved that consumers with higher fashion involvement are likely to use new technology, however are unconvinced to buy an AI integrated fashion product. 3. Consumers responded positively to the functions and user friendly system of the AI product which aided them in selecting trendy outfits. Hence, resulting in positive attitudes towards AI.
2. Rathore, 2017	Impact of using AI and the metaverse in fashion marketing	Qualitative	Developing countries (India)	1. Consumers digital literacy and the user interface 2. Technological issues 3. Brand innovation 4. Marketing effort	1. The usage of AI improves understanding of consumer behavior, engagement and brand innovation. Resulting in better outcomes of metaverse marketing efforts. 2. AI technology in fashion aids in providing hyper personalized shopping experiences, tracking and forecasting trends, enhancing consumer interaction through understanding emotions and creating body inclusive designs. 3. AI-powered chatbots, virtual reality shopping experiences, or AI-driven predictive trends directly impact the

					creation of interactive and personalized experiences in the virtual world. 4. Some major threats of integrating AI in fashion marketing include digital literacy, user interface, technological issues, data privacy, legal constraints etc.
3. Brüns and Meißner, 2024	Consumer perception of brands which use GenAI for social media content creation	Quantitative and Qualitative	Developed countries (Germany)	1. Perceived brand authenticity 2. Algorithm aversion	1. Adopting GenAI gives rise to negative attitudes and behavioral reactions among followers / consumers of brands. However, these after effects are also mediated by followers' perceptions of brand authenticity. 2. Negative reactions can be reduced if GenAI is used to aid humans with content creation as opposed to automation replacing them entirely. 3. To benefit from economic rewards whilst also maintaining positive relationships with consumers, brands require to subtly change their way of GenAI adoption.
4. Xu and Mehta, 2022	Consumer response to AI-designed luxury products	Quantitative	Developed countries (USA)	1. Type of luxury product 2. Perceived significance of products emotional and functional value 3. Perceived brand essence 4. Consumer purchase intentions	1. Incorporating technology in luxury product design unfavorably impacts the emotional value yet improves the associated functional value of the product. This causes a varied effect on consumer patterns and factors like perceived brand essence, brand attitude and consumer purchase intentions. 2. Using AI in designing luxury products reliant on superior emotional value (e.g., fashion brands), diminishes perceived brand essence and produces negative consumer reactions. 3. However, the negative response is lesser if the brand's essence also includes functional value (e.g., luxury cars) or this value is highlighted, for example through marketing.
5. Abdul et al., 2021	Relationship between branded content and social media	Quantitative	Developed countries (Spain)	1. Specific brand 2. Brand content category (eg content including social responsibility)	1. Social networks are found to make interaction between brands and their followers easier. They allow content and messages to have a higher impact on the

	consumer interactions of the luxury brand Manolo Blahnik				public, hence increasing the brand's engagement with its followers. 2. Manolo Blahnik's followers responded positively to brand content published on facebook which provoked optimistic reactions. 3. Brand content influences the sharing of posts, feelings towards the brand and its positive diffusion amongst people.
6. Rahman et al., 2023	Effects of AI-powered digital assistance towards consumers' luxury brand online shopping experience	Quantitative and Qualitative	Developing countries (Oman)	1. Customer engagement 2. Online shopping experience 3. Personalisation 4. Competitive	1. Popular luxury brands (eg: Louis Vuitton, Burberry, Dior, Tommy Hilfiger, Estee Lauder, Sephora, Farfetch, L'Occitane and Neiman Marcus) are beginning to use AIPD (AI-powered digital assistance) and DMC (digital multisensory cues) applications to enhance their consumers' online shopping experience. 2. To influence customer engagement AIPDA is one of the prime antecedents. 3. Luxury brands should make use of AIPDA within their online platforms as it leads to more personalisation for consumers. 4. Luxury brands should surely invest in new technologies such as AI for survival and remaining competitive.
7. Lee and Kim, 2023	Consumer response to AI designed fashion products, focusing on clothing	Quantitative and Qualitative	Developed countries (USA)	1. Perceived authenticity of clothing 2. Consumers' preexisting fashion design schema	1. Findings overall suggest consumers respond negatively to AI-designed clothing, compared to human-designed. Caused due to, the perceived authenticity of AI being low as a result of the differences between consumers' preexisting fashion design schema and AI generated ones. 2. Study 1 & 2 (282 consumers) indicated that the response toward human designed clothing to be more favorable as its perceived authenticity and expected product quality is higher. 3. Study 3 (156 consumers) showed that offering customization options for AI-designed clothing improves consumer attitudes and purchase

					intentions. Therefore, attitudes towards AI-customisation are more positive compared to non-AI-customisation.
8. Kumar et al., 2024	Forecasting consumer buying behavior of luxury goods using machine learning algorithms	Qualitative	World	1. Competitive landscape 2. Changing trends 3. Predictive analysis 4. Marketing strategy 5. Inventory management	1. To flourish in the competitive landscape of luxury fashion in this digital age, it is vital to make use of machine learning algorithms. 2. Through the power of predictive analysis, businesses can improve their understanding of consumer preferences while simultaneously being a step ahead of changing trends. 3. Embracing machine learning can have visible advantages for marketing strategy, inventory management, and product development in luxury fashion.

Content Analysis

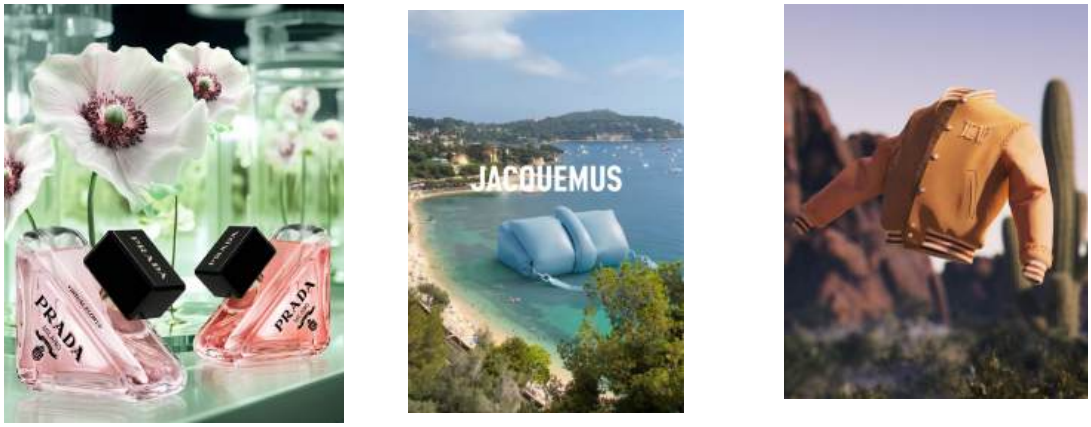
The evidence collected from the social media marketing and websites of luxury fashion brands provides key insights for the study. Artificial Intelligence (AI) is being utilized across various industries, like the luxury fashion industry, for ideation, designing, digital assistance and more. However, the dilemma of integrating AI into generating marketing content for luxury fashion products, and its impact on consumer purchase intention has yet to be thoroughly explored (Xu and Mehta, 2022). The following content analysis offers assistance to luxury fashion companies navigating the world of artificial intelligence in relation to luxury fashion marketing. It examines how luxury fashion brands utilize AI-generated marketing content to influence consumer purchase intentions. Themes observed within the content analysis are further validated by interviews conducted with experienced luxury product buyers. The detailed findings are shared below. First, the use of an analogous color-based theme amongst AI generated marketing content was a common pattern observed. Colors located next to each other on the color wheel, such as pink, purple and blue are examples of analogous colors. Social media posts advertising luxury fashion products such as the Moncler Jacket (Figure 1) incorporated varied shades of blue hues in the posts' background to enhance its overall visual impact on consumers. Blue-green analogous colours tend to elicit positive attitudes among consumers, driving purchase intention (White et al., 2021). AI-generated product advertisements for Louis Vuitton follow the same pattern, displaying bright analogous colors (orange, yellow, brown) in Figure 1. The analogous color scheme made use of in the background, supports thematic consistency while subtly focusing consumer attention onto the product being marketed. The harmonious color scheme seamlessly integrates the product into the visual narrative, enhancing its attractiveness.

Figure 1 : Color-based theme



Second, the natural background integration of AI-generated marketing content often featured natural settings, abstract elements or landscapes that complimented the products color. Incorporating elements such as flowers and oceans in the background, support thematic consistency of the post and blend in with the product color. As observed in Figure 2, the flower placement in the background compliments the Prada perfume being marketed and hints at its floral and sweet fragrance supporting thematic consistency. Featuring natural elements and landscapes focuses consumer attention onto the product without striking backgrounds or colors. Furthermore, creating visual harmony to amplify the product's appeal to the human eye.

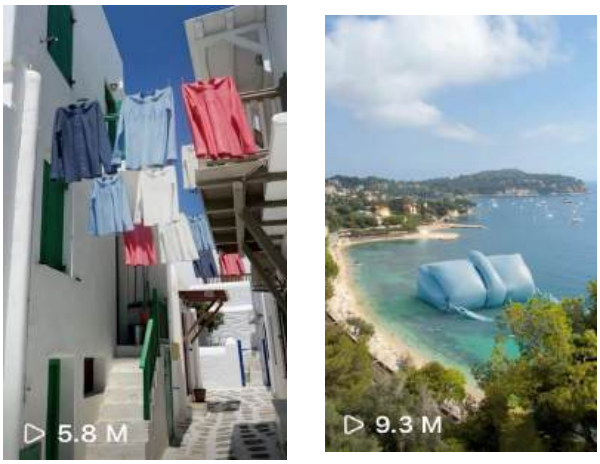
Figure 2 : Natural background integration



Next, engagement metrics and emotional responses to luxury brands' social media posts incorporating AI-generated content were higher and more positive than for their regular posts. On average, brands in the luxury fashion industry, such as Tommy Hilfiger (Figure 3), receive around 1 million views per Reel and anywhere between 10,000 to 50,000 likes, varying based on the content (Landry, 2024). As exemplified in Figure 3, a Reel containing AI-generated content created by Tommy Hilfiger advertising its shirts gained 5.8 million views and 47,000 likes, suggesting the more attractive nature of the content. Brands such as Jacquemus also saw higher

interaction rates and views, with an AI-generated Reel garnering 9.3 million views and 530,000 likes, compared to their average of 1 to 3.5 million views and 50,000 to 150,000 likes per Reel. Upon conducting a sentiment analysis on the comments for these posts, fascination and amazement at the creativity and uniqueness of the AI-generated content were among the most observed emotions of viewers. The comment sections of these posts were generally positive. The frequent use of positively connotated emojis also signified an optimistic response toward AI-generated marketing content.

Figure 3 : High engagement metrics



Fourth, a common finding in AI-generated marketing content is the use of human mimicking AI models. AI creates models in marketing ads that appear realistic with a slightly animated touch to differentiate them, thereby creating a sense of relatability between the marketing ads and consumers. Additionally, this approach allows luxury fashion brands like Valentino and Levi's (Figure 4) to incorporate uniqueness among their models, setting them apart from competitors. Besides using life-like AI-generated models, Tommy Hilfiger ensured the models appeared relatable to their target audience by launching 'Noah,' an ultra-realistic digital idol in a virtual space in China. Tommy Hilfiger localized Noah's appearance to symbolize Chinese features, making him widely popular and gaining over 400,000 followers (Figure 4).

Figure 4 : Human mimicking AI models



Lastly, AI-generated content aligns well with the brand identity. The storytelling within the content conveys a uniform perception and identity of a brand. This is exemplified by Coperni, a brand known for its unique approach to marketing. Coperni's ethos relates to a technology-oriented approach to innovation in the field of fashion. Figure 5 represents AI-generated visuals of their AW23 digital campaign showcasing the interaction between a human and a robot in abstract natural landscapes. The use of darker, minimalistic colors aligns with the brand's identity and showcases their bold, distinctive luxury status. AI-generated content further enhances consumer perception of the brand, portraying their modern and adaptive nature toward technological trends. This emphasizes a futuristic approach to fashion.

Figure 5 : Alignment with Brand Identity



Interviews

Interviews were conducted with the target demographic of Indian adults who have experience purchasing three or more luxury fashion products in the past and are well acquainted with technology. Conducting interviews allows an in-depth exploration of participants' attitudes and perceptions, providing a comprehensive understanding of how AI-generated marketing can influence their purchase intentions. Eight adults, both men and women between the ages of 30 and 50, were interviewed. The interviews were conducted either through real-life conversations or the online platform Zoom, both of which ensured participants could see the interviewer and clear any doubts face-to-face, making it a comfortable experience.

Participants were given the choice to select from a wide range of interview timings spanning from 14th to 21st July, ensuring sufficient time for data collection. The participant selection process incorporated individuals with different careers, lifestyles, and areas of interest, providing diverse viewpoints. Participants were mainly asked open-ended questions along with a few close-ended questions. Visual aids such as photographs and videos were provided as examples to equip participants with a better understanding of AI-generated marketing content.

The visuals shown for comparison between AI generated and human generated content for marketing involves all categories and styles of fashion apparel to make it inclusive and ensure that we validate the findings from the interviews. Topics included how important marketing content is for them when making luxury fashion purchases and whether they would prefer AI-generated marketing content over regular marketing content based on examples. Broader ideas covered their experience with luxury fashion marketing, opinions on the integration of AI, how AI marketing influences their opinion of a brand, and more. For a detailed overview of the interview questions, structure, and visual aids used, please refer to Appendix A, Table 2.

Responses and insights from the interviews indicated a balance between those in favor of AI-generated marketing and those who are not. A common response among participants was that AI marketing is far superior to traditional marketing in terms of innovativeness, but the reverse is true for effectiveness. Most interviewees suggested that AI integration in marketing can increase brand value, reflecting the positive responses toward AI-generated marketing. Opposingly, there were adverse reactions to AI replacing humans in marketing, with many participants stating that AI marketing can appear unconvincing, artificial, and may inaccurately portray the luxury product. When asked to choose between traditional marketing posts and AI-generated marketing posts for different luxury products, the overall response was equally divided. Additionally, more consumer experience was inversely related to the influence of AI marketing on purchase intention for luxury products. Lastly, the lack of familiarity with AI and the incorrect targeting of AI-generated luxury marketing may also contribute to the negative response. The overall response was largely equally divided, including both praise and criticism of AI in luxury marketing.

Discussion

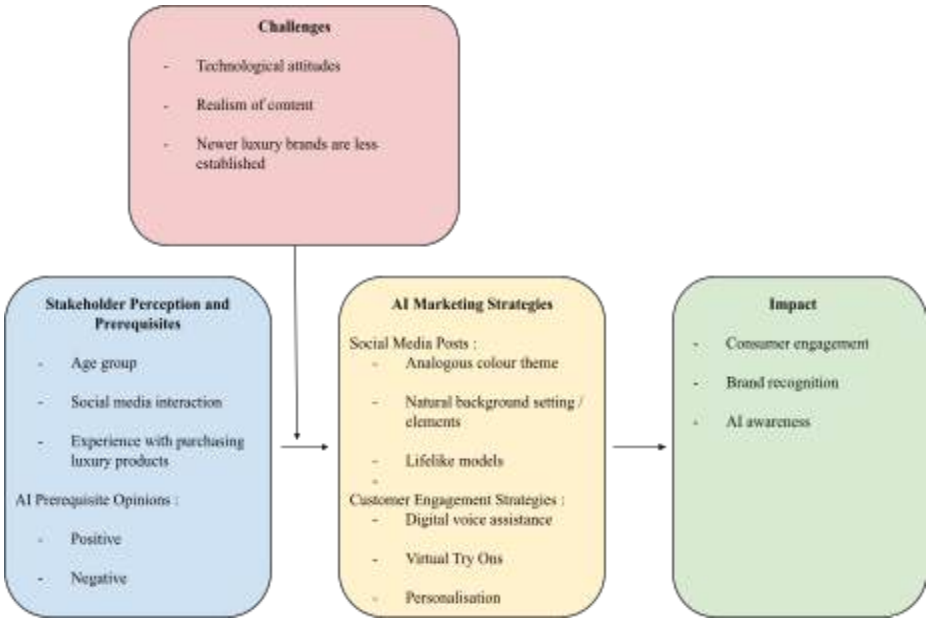
After conducting various forms of analysis and data collection, we gained a nuanced understanding and multiple perspectives on consumers' views regarding AI-generated luxury marketing. The insights gathered from the extensive literature review, content analysis, and interviews offered a valuable understanding of how AI-generated luxury marketing can impact consumer responses and purchase intentions toward products.

Observations from the literature review and interviews were initially compiled and categorized into subcategories based on key ideas and terminology used. AI marketing strategies were gathered via the content analysis. Next, all findings were consolidated, and broader themes amongst them were identified. Based on each theme, separate boxes were created in the framework. Overall, the framework clearly illustrates the flow of themes across our findings and comprehensively addresses each aspect of the research question, including AI-generated marketing, consumer impact, and purchase intention.

First, stakeholder perceptions and prerequisites are introduced, exploring key factors that influence a consumer's relationship and response to AI-generated luxury marketing on social media. Herein, we also explore the positive and negative consumer perceptions gathered from our interviews. Next, AI marketing strategies used on Instagram, as discussed in the content

analysis and literature review, are highlighted. Possible challenges faced by luxury fashion companies when using AI in marketing are identified in connection to stakeholder perceptions and AI marketing strategies. Lastly, the notable impact of these AI marketing strategies is analyzed and developed. The detailed framework is recorded below.

Framework Diagram



Challenges

This component of the framework identifies obstacles faced by luxury fashion companies when attempting to use AI generated marketing content to promote their products on social media platforms. Some challenges may include Technological attitudes, Realism of content and Newer brand being less established.

According to Forbes India (Afprelaxnews, 2022), existing luxury product buyers in the country comprise of mainly Millennials between the ages of 28-43 years. Some of these consumers are not well accustomed to 21st century technology and social media platforms such as Artificial Intelligence and Instagram. Hence, as stated by Rathore, 2017 “AI might be intimidating for traditional users" in fashion marketing due to the field's unexplored nature. Moreover, they may also have a subconscious bias towards “old school” marketing content. resulting in consumer openness and reactions towards posts which incorporate AI generated content may be negative as they would prefer “more classic” marketing campaigns. Bridging this digital divide between consumers and AI marketing can prove to be a roadblock for luxury fashion companies.

Several interview participants stated the “unrealistic”, “gimmicky” and “unrelatable” nature of AI generated marketing content from luxury brands such as Louis Vuitton and Tommy Hilfiger. Rathore, 2017 elaborates on this in relation to achieving true-to-life realism digital

representations of clothing, texture and fitting is not straightforward. This suggests that luxury brands should look at creating more realistic and relatable AI generated content, which consumers find appealing and convincing of purchasing the product.

In addition, recently launched luxury fashion brands would find it hard to compete with household names of luxury fashion houses which are famous across the globe. Newer brands are unlikely to have the financial means or resources to spend on AI generated marketing, therefore would be unable to compete with established luxury brands. Such companies also rely heavily on creating extensive marketing content to attract a broad consumer base, whereas prominent luxury brands are already well-established and do not depend primarily on marketing for recognition.

Stakeholder Perception and Prerequisites

This researcher found consumers in the luxury fashion industry to be the main stakeholders. Certain perceptions and prerequisites amongst them can prove to influence their purchase intention towards luxury products upon seeing AI generated marketing content.

Based on the consumer age group of brands their purchase intention after interacting with AI generated content gets influenced. The integration of AI in marketing appeals more to the younger generation of consumers, thereby fueling their purchase intention towards the product (Gupta et al., 2024). As observed in our interviews as well, participants within the age of 30-35 were on average likely to be more influenced by AI generated content unlike those over the age of 35.

Next, the social media interaction levels of consumers can prove to be an important prerequisite as they convey whether consumers are accustomed to engaging with marketing content. Being highly active on social media and encountering more luxury product AI generated posts/advertisements could cause them to get persuaded and influenced to purchase the product online.

In addition, the more experience a consumer has in purchasing luxury fashion products, the less likely they tend to get influenced by marketing content whether it is traditional marketing or AI. People who are new at buying luxury products typically lack the knowledge and decision making ability to make an independent firm choice, hence are influenced by marketing and posts on platforms such as Instagram. Whereas an experienced luxury buyer "knows what they want to purchase based on utility, choice, and value for money" as observed within our interviews.

AI Prerequisite Opinions

1.Positive

Firstly, all interview participants suggested that AI marketing designs and posts tend to be extremely innovative and "far more graphically appealing". Participant 2 stated that AI marketing would influence a positive purchase intention towards luxury fashion products as its portrayal of brand value surpasses traditional marketing. Thirdly, multiple participants' responses toward AI marketing strategies such as Virtual Try Ons, Digital Voice Assistance and

Personalisation were immensely positive. Such AI marketing strategies garnered an optimistic response amongst these consumers as they increased brand value, offered assistance and acted as a brand USP.

2.Negative

The participant response to AI marketing did have negative aspects too. In many statements, the ineffectiveness of AI marketing was criticized by participants. Emphasis was placed on posts' unrealistic appearance and their incorrect target audience in the Indian population. Human touch in luxury marketing seemed to be favored due to the unfamiliarity of and lack of knowledge about AI marketing. Participants clearly stated that human touch in marketing cannot be replaced or taken over by AI.

AI Marketing Strategies

Social Media Posts

One of the vital visual elements used in luxury brands' social media posts of AI generated content was an analogous color scheme. Using complementary colors in such posts can make the product look more attractive and ensure a subtle focus is placed on it. People may find this type of AI generated marketing content more appealing and harmonious to look at. Furthermore, multiple brands have utilized natural elements and background settings to complement product colors in their AI-generated marketing content. Posts featuring landscapes such as deserts, Greek city streets, and oceans have garnered significant interest among viewers. These settings support the thematic consistency of the posts, making them more convincing and appealing for potential buyers. Additionally, another marketing strategy seen amongst AI generated content on social media posts is the use of life-like models. By creating models which appear realistic with a slightly animated look to them, luxury brands are showcasing a sense of relatability to their consumers. The use of these life-like models has gained popularity among consumers and has become an effective tool for brands to influence consumer purchasing decisions. By viewing luxury products on these AI models, many consumers may be persuaded to purchase the products in reality.

Consumer Engagement Strategies

Beginning in 2023, popular luxury brands such as Burberry, Dior, Tommy Hilfiger have started making use of AIPDA (AI-powered digital assistance) and DMC (digital multisensory cues) applications to enhance their consumers' online shopping experience (Rahman et al., 2023). AI marketing applications like these lead to increased personalization of product ranges and advertisements for consumers. By doing so, consumers are shown products tailored specifically to their interests, resulting in a more focused selection. This targeted approach can improve consumers' shopping experience along with increasing the chances of purchases, since consumers are likely to see items which align with their preferences. All in all, the use of AI in

marketing in the form of Virtual try-ons, Digital Voice Assistance, and Personalisation are a unique selling point for brands. They offer solutions and assistance to consumers whilst shopping therefore increase a luxury brand's value from a consumer viewpoint resulting in possibly increased purchases.

Impact

The impact component of this paper offers an in-depth evaluation of the AI marketing strategies used by luxury fashion brands, considering stakeholder perceptions and challenges encountered along the way. Impacts may include, Consumer Engagement, Brand Recognition and AI awareness.

First, the use of AI-generated marketing content on luxury brands' social media pages has a profound impact on consumer engagement. Metrics such as likes and views on posts containing AI-generated content are noticeably higher than those without. The innovative AI visuals increase consumer connection and intrigue with brand posts, broadening the reach of luxury brand advertisements and attracting new consumers. Engaging audiences with compelling AI content heightens engagement metrics, potentially leading to positive purchase intentions.

Additionally, luxury brands gain increased recognition from audiences when they utilize AI-generated marketing. The unique visuals and creative elements serve as differentiators in luxury marketing, fulfilling the role of a USP. This approach demands more consumer attention, making the content memorable and increasing the brand's chances of capturing consumer attention and convincing them to make a purchase.

Finally, awareness of AI grows among consumers. By interacting with various AI marketing elements and posts online, audiences become more knowledgeable about AI technology. This increased exposure educates consumers about AI's capabilities while simultaneously leading them to accept and trust AI-driven content more. As consumers become increasingly familiar with AI, they are more likely to engage with it and develop a positive outlook on its applications in luxury marketing, driving consumer purchase intention and interest in the field.

Implications

This study aims to provide insight into the unexplored impact of AI marketing in luxury fashion on consumer purchase intention. The significant implications of this study address a multitude of stakeholder groups, including potential customers, marketers, luxury fashion companies, and technologists. While most previous papers primarily discuss the clothing industry, luxury fashion is rarely specified as their focus. Additionally, although prior studies have covered the impact of AI in fashion and product design, none have delved into its role in marketing. Hence, this study addresses a gap in past research that needs to be overcome. Furthermore, most past studies are based in developed nations like the USA, with few focusing on developing countries such as India. In recent years, India's luxury market has witnessed significant growth, largely propelled by its youth demographic. The number of luxury consumers

in the country is estimated to rise from approximately 60 million in 2023 to 100 million by 2027 (Ksheerasagar, 2024), making the Indian demographic important to explore.

Research on this topic is useful, as investigating the effect of AI-generated marketing content on consumer loyalty and behavior can greatly aid luxury retail companies and other business owners in making informed decisions about integrating AI into their marketing strategies. The luxury fashion industry is projected to grow substantially, becoming more competitive. To distinguish their brand from others, it is necessary for a luxury fashion company to create unique marketing campaigns, with AI being a prime example of how to do so. Additionally, marketing holds immense importance in influencing purchases. Incorporating attractive and appealing visual marketing elements is imperative in increasing consumer engagement and encouraging the purchase of a luxury brand's product. Therefore, understanding the influence of AI on brand marketing can be crucial for companies' sales, influence, and growth across the globe.

Limitations

While conducting research, it is inevitable for a few limitations and obstacles to arise. Firstly, due to the author's limited network, the number of interview participants is relatively small. The small sample size reduced the scope of perspectives from interviews. However, since most participants were highly experienced luxury product buyers, their responses did offer a diverse array of viewpoints. In the future, researchers with a broader network of participants can discuss this topic in greater detail.

Secondly, the research was conducted with a limited geographical focus on India, specifically the city of Mumbai. As of 2023, Mumbai’s average income is among the highest in the country, and the city is known as one of the most luxurious in India. Therefore, the data and results may be considered biased and may not be applicable to other regions globally due to social and cultural discrepancies. Future research should account for this limitation and explore a wider geographical focus to ensure accuracy.

As most of the data collected is qualitative, cognitive bias may occur, which should be taken into account. Interview participants have varying financial values and spending patterns regarding luxury purchases. Based on their principles and susceptibility to marketing influence, each may be subconsciously biased. Acknowledging this limitation and offering alternatives to conduct experiments that avoid subconscious bias can aid future research. Additionally, the short time period assigned for data collection could be the cause of the limited number of responses. To resolve the challenge of limited responses, future research should aim to collect data over a less restrictive time period.

Lastly, both the luxury fashion and technological industries are continuously evolving, with new developments being made daily in relation to Artificial Intelligence. Thus, the findings and information from this study may become slightly outdated in the future. However, this does allow future researchers to delve into this topic in relation to more advanced versions of AI and evolved marketing strategies.

Conclusion

In conclusion, this study aimed to delve into the growing influence of AI-driven marketing in the luxury fashion industry while addressing gaps in previous research. Through a comprehensive mix of methods, using literature review, content analysis and in depth interviews for data collection this research sought to examine the impact of AI generated marketing content on consumer purchase intention towards luxury fashion products.

The findings from the content analysis of luxury brands' AI-generated social media posts highlight recurring themes. These themes include the use of analogous color palettes, life-like AI models, alignment with brand identity, and natural background elements. Overall, these themes display a positive outlook and responses toward AI-generated luxury marketing content, further supported by high engagement metrics. This demonstrates that the use of AI in social media can be unique, innovative, and beneficial for a brand, positively influencing consumer purchase intentions.

Interviews provided a diverse set of responses and opinions on the topic. Reactions toward AI-generated visuals of various luxury products, from clothing to handbags, were evenly split between AI marketing content and traditional marketing content. For each product visual, 50% of participants preferred the AI-generated version, while the other 50% chose the traditional one. Responses indicated that AI is highly praised for its innovativeness and problem-solving capabilities, but it is not viewed as effective in convincing customers to purchase compared to traditional marketing. Participants suggested that AI marketing should target the younger generation more and noted that Indian consumers need more education about AI to be convinced to purchase expensive luxury products. As AI evolves, its effectiveness and impact on consumer trust and reliability may improve.

This study concludes that AI should be used in visual marketing on social media posts, but in a limited manner. While AI-generated marketing content positively influences consumer purchase intention, this effect occurs only under certain constraints. First, AI visuals should be realistic yet innovative, providing an accurate depiction of the luxury product being sold rather than appearing fictitious. Additionally, luxury fashion companies should target the younger Indian demographic due to their higher technology usage and awareness. The general consensus also indicates that human involvement is still necessary in luxury marketing and cannot be entirely replaced by AI. Overall, the findings from this study should guide luxury fashion and marketing companies, as well as related individuals, in their future strategies to ensure success in the field.

Works Cited

Afprelaxnews. “Millennials and Gen Z Are Now More Likely to Embrace Luxury than Their Elders.” *Forbes India*, ForbesIndia, www.forbesindia.com/article/lifes/millennials-and-gen-z-are-now-more-likely-to-embrace-luxury-than-their-elders/73231/1. Accessed 21 Aug. 2024.

Bhandari, Pritha. “Ethical Considerations in Research: Types & Examples.” *Scribbr*, 9 May 2024, www.scribbr.com/methodology/research-ethics/.

Brüns, Jasper David, and Martin Meißner. "Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity." *Journal of Retailing and Consumer Services* 79 (2024): 103790.

Castillo-Abdul, Bárbara, Mónica Bonilla-del-Río, and Estela Núñez-Barriopedro. "Influence and relationship between branded content and the social media consumer interactions of the luxury fashion brand Manolo Blahnik." *Publications* 9.1 (2021): 10.

Du, Duo, et al. “Effect of AI Generated Content Advertising on Consumer Engagement.” *SpringerLink*, Springer Nature Switzerland, 1 Jan. 1970, link.springer.com/chapter/10.1007/978-3-031-36049-7_9.

Gupta, Vandana, Junaid Hushain, and Abhilasha Mathur. "The Future of Luxury Brand Management: A Study on the Impact of New Technology and Relationship Marketing." *AI in Business: Opportunities and Limitations: Volume 1*. Cham: Springer Nature Switzerland, 2024. 57-68.

Ksheerasagar, A. “Luxury Revolution Sweeping India: What’s Fueling the Shift in Consumer Preferences?: Stock Market News.” *Live Mint*, 1 Apr. 2024, www.livemint.com/market/stock-market-news/luxury-revolution-sweeping-india-whats-fueling-shift-in-consumer-preferences-louis-vuitton-gucci-burberry-rolex-11711939444128.html.

Konina, Natalia Yu. "Artificial Intelligence in the Fashion Industry—Reality and Prospects." *Anti-Crisis Approach to the Provision of the Environmental Sustainability of Economy*. Singapore: Springer Nature Singapore, 2023. 273-280.

Kumar, Anil, et al. "Combined artificial bee colony algorithm and machine learning techniques for prediction of online consumer repurchase intention." *Neural computing and Applications* 31 (2019): 877-890.

Lamont-Djite, Tara. “Brand Watch: Coperni.” *Harpers Baazar* , 8 Apr. 2015, www.harpersbazaar.com/fashion/fashion-week/a10243/brand-watch-coperni/.

Landry, Jamie. “Instagram Reels Stats and Benchmarks: Dash Hudson.” *Instagram Reels Stats and Benchmarks | Dash Hudson*, Dash Hudson, 11 July 2024, www.dashhudson.com/blog/instagram-reels-performance-benchmarks.

Lee, Garim, and Hye-Young Kim. "Human vs. AI: The battle for authenticity in fashion design and consumer response." *Journal of Retailing and Consumer Services* 77 (2024): 103690.

Liang, Yuli, Seung-Hee Lee, and Jane E. Workman. "Implementation of artificial intelligence in fashion: are consumers ready?." *Clothing and Textiles Research Journal*38.1 (2020): 3-18.

“Mumbai Retains Top Spot as India’s Most Expensive City for Expats: Mercer’s 2023 Cost of Living Survey.” *Mumbai Retains Top Spot as India’s Most Expensive City for Expats: Mercer’s 2023 Cost of Living Survey*, www.mercer.com/en-in/about/newsroom/mumbai-retains-top-spot-as-indias-most-expensive-city-for-expats/. Accessed 4 Aug. 2024.

Rahman, Muhammad Sabbir, et al. "The new wave of AI-powered luxury brands online shopping experience: The role of digital multisensory cues and customers’ engagement." *Journal of Retailing and Consumer Services* 72 (2023): 103273.

Rathore, Bharati. "Exploring the intersection of fashion marketing in the metaverse: Leveraging artificial intelligence for consumer engagement and brand innovation." *International Journal of New Media Studies: International Peer Reviewed Scholarly Indexed Journal* 4.2 (2017):51-60.

“The Coperni-Can Revolution: What’s behind the 99% Air Swipe Bag?: Istituto Marangoni.” *ISTITUTOMARANGONI*, Accessed 29 July 2024. www.istitutomarangoni.com/en/maze35/game-changers/the-coperni-can-revolution-whats-behind-the-99-air-swipe-bag.

Tommy, Hilfiger. “Creating a ‘Realistic’ Virtual Campaign with Tommy Hilfiger.” *Wavemaker Global*, 1 Aug. 2023, wavemakerglobal.com/our-work/growth-stories/noah-x-tommy-hilfiger.

White, Allison-R., et al. “Color in Web Banner Advertising: The Influence of Analogous and Complementary Colors on Attitude and Purchase Intention.” *Electronic Commerce Research and Applications*, Elsevier, 18 Oct. 2021, www.sciencedirect.com/science/article/abs/pii/S1567422321000727.

Xu, Lidan, and Ravi Mehta. "Technology devalues luxury? Exploring consumer responses to AI-designed luxury products." *Journal of the Academy of Marketing Science* 50.6 (2022): 1135-1152

The Molecular Landscape of Small Lymphocytic Lymphoma: Mechanisms and Therapeutic Approaches By Sam Mahdavy

Abstract

Cancers such as Small Lymphocytic Lymphoma (SLL) and Chronic Lymphocytic Leukemia (CLL) are typically caused by a series of mutations within an individual's genome. Genes such as *TP53*, *BCL-2*, and *MYC* are crucial in the progression of not only these diseases, but others too. Therefore, in developing therapies for these disorders, it is essential to understand the mutations and interactions of these genes at the molecular level. As a result, understanding the complex mechanisms and signaling pathways that connect the genes is paramount to developing treatments. Additionally, when investigating a cancer case, the pathology behind these cancers must be accurately interpreted for a proper diagnosis. Cytokines must also be understood as they are involved in the cell signaling process. Regarding treatments, comparing the effects of different drug combinations is necessary to determine the most effective treatment regimen. In addition to drug-based treatments, immunotherapies also prove effective. Despite these advanced technologies, diagnosing cancer early on is essential, as an early diagnosis significantly improves the chances of successful treatment and survival.

Introduction

Small Lymphocytic Lymphoma (SLL) is a cancer of uncontrolled, proliferating mature B lymphocytes in the lymph nodes. T lymphocytes are also involved but are not the main source of cancerous cells. Additionally, Chronic Lymphocytic Leukemia (CLL) is the same disease as SLL. However, most cancerous cells in CLL are located in the bloodstream and bone marrow. SLL and CLL are examples of Non-Hodgkin Lymphomas (NHLs) that begin in the lymphocytes, or white blood cells (WBCs) that facilitate the body's response to infection. SLL is staged using the Lugano classification system, which takes into account the extent to which the cancer has spread within the lymph system (Table 1).

Stage	Diagnostic criteria
Stage I	SLL found in one lymph node or a group of adjacent lymph nodes
Stage II	Found in at least two groups of lymph nodes on same side of the diaphragm
Stage II	Same as above with "bulky" or tumor of large size
Stage III	Found in areas of lymph node above and below the diaphragm and in the spleen
Stage IV	SLL spread widely to at least one organ outside the lymph system

Table 1: Lugano Classification System (adapted from Yoo 75 - 8)

Conversely, CLL is typically classified with the Rai staging system based on blood tests (Table 2).

Stage	Diagnostic criteria
Rai stage 0	High lymphocyte count (lymphocytosis); no enlargement of lymph nodes, spleen, or liver; red blood cell (RBC) and platelet counts are approximately normal
Rai stage I	Lymphocytosis; enlarged lymph nodes; no enlargement of spleen or liver; potentially slightly low RBC and platelet counts
Rai stage II	Lymphocytosis; possible enlarged lymph nodes; enlarged spleen; possibly enlarged liver; potentially slightly low RBC and platelet counts
Rai stage III	Lymphocytosis; RBC deficiency (anemia); possible enlarged lymph nodes, spleen, and liver; approximately normal platelet counts
Rai stage IV	Lymphocytosis; platelet deficiency (thrombocytopenia); possible anemia; enlarged lymph nodes, spleen, or liver

Table 2: Rai Staging System (adapted from Rai et al. 219 - 34)

In this review, we will explore some of the genes involved in these cancers as well as the mechanisms by which they progress. Furthermore, we will discuss the structure and knockout mouse specific to each gene to understand the function of the gene itself in a broader context.

TP53: The Critical Axis and Normal Function

TP53 is a gene that encodes the tumor suppressor protein p53. It suppresses tumor development by regulating cell division, prohibiting excessive and uncontrollable cell growth. p53 is located in the cell nucleus and binds to DNA. At baseline in a healthy cell, p53 levels are kept low by the negative regulatory proteins MDM2 and MDMX, which mark p53 protein for degradation via a process called ubiquitination. However, when DNA becomes damaged by radiation, UV light, or a lack of nutrients or oxygen, p53 expression increases to determine whether the replicated DNA will be repaired or the cell will undergo apoptosis. Specifically, ubiquitination is inhibited, leading to a rapid growth of intracellular p53 levels. Then, these proteins are activated and stabilized through processes such as phosphorylation, acetylation, and methylation (Kubbutat et al. 299 - 303). If the DNA can be repaired, p53 activates specific pathways to facilitate this process. If not, stabilized p53 protein creates tetramers in the nucleus and binds to DNA to regulate gene transcription, leading to apoptosis (Williams and Schumacher 26070). Through this process, p53 prevents tumor development by regulating the quality of replicated DNA. However, if the *TP53* gene is mutated, the protein's function is compromised, and mutant DNA might continue to cause cancers such as SLL due to the absence of the regulatory protein.

p53: Structure

The structure of p53 consists of 393 amino acids including the following: an N-terminal transactivation domain (TAD), a proline-rich domain (PRD), a DNA binding domain (DBD) connected to a tetramerization domain (TET), and a C-terminal regulatory domain (CTD). The TAD contains an acidic and intrinsically disordered TAD and a proline-rich area. TADs like TAD 1 and TAD 2, both of which are contained in p53, bind to transcriptional machinery components and coactivators to promote transcriptional initiation and interact with negative regulators to end transcriptional activation (Kastenhuber and Lowe 1062 - 78). TAD can also adopt an amphipathic α-helical conformation when binding to partner proteins, such as MDMX and MDM2. The PRD contains five PXXP motifs, some of which may have a polyproline helix-like structure. However, the PRD's exact structure is unknown. The DBD structure contains an immunoglobulin-like β-sandwich, a loop-sheet-helix motif and two large loops stabilized by a tetrahedrally coordinated zinc ion. Lastly, the CTD is an intrinsically disordered region, similar to the TAD, with positively charged residues and a zinc-binding domain (Wang et al. 92).

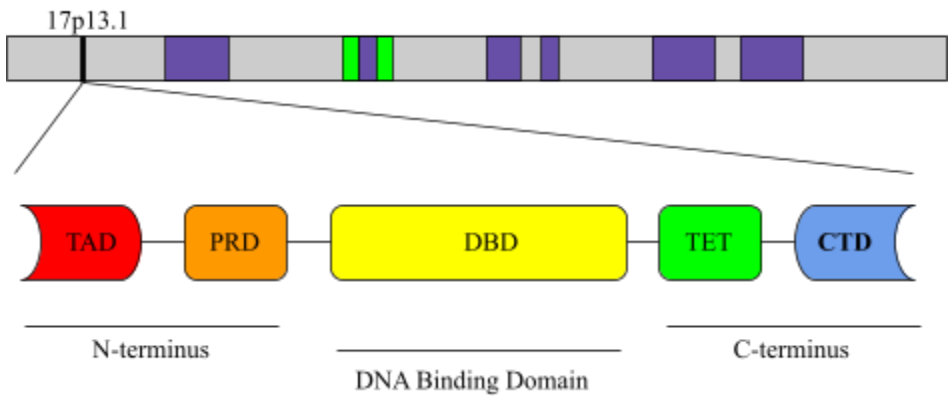


Fig. 1: Domain architecture schematic for p53 (adapted from Babamohamadi et al. 903075)

TP53: Knockout Mouse

Several research groups have investigated the impact of deactivating the *TP53* gene in mice. Knockout mice with a deactivated *TP53* gene were found on average to develop a tumor in 4.5 months and later die within the next 1.5 months. In contrast, half of the control mice, which were heterozygous, developed tumors within 18 months and were able to live for an additional 12 months (Donehower 269 - 78). Therefore, the inhibition of the *TP53* gene demonstrates that in the absence of a tumor suppressor, a tumor can develop four times as quickly. Furthermore, this highlights why tumors must be diagnosed early to prevent damage to an individual's health.

TP53: MicroRNAs and Mutations

Typically, the cause of cancer is linked to alterations in the genetic code of genes such as *TP53* at the molecular level. The molecular mechanisms underlying p53 and cancer development

involve microRNA (miRNA) molecules and their interactions with T lymphocytes. A microRNA is an endogenous single-stranded RNA molecule that binds to complementary sequences of target mRNAs. This process leads to mRNA suppression of translation. For example, miR-34a is a miRNA that is transactivated by p53 targeting oncogenes FOXP1 and BCL-2. The *TP53* gene induces expression of miR-34a through direct binding to its promoter. When miR-34a is underexpressed, the results include diffuse large B-cell lymphomas (Xu-Monette et al. 3668 - 83). Another miRNA is miR-21 which is regulated by the transcription factor FOXP3 which governs the function of T regulatory cells. Also, miR-21 displays anti-apoptotic properties in malignant T cells. Specifically, miR-21 mutations prevent it from inhibiting the development of cancer which leads to the inability to conduct apoptosis. However, it is currently unknown which specific cells express and which mechanisms regulate miR-21. Similar to the effects observed with miR-21, mutations in miR-155 yield comparable outcomes. miR-155 is expressed when the protein STAT5 is activated, which contributes to the proliferation of malignant T cells (Lindahl et al. 45730 - 44). In total, these pathways contribute to the proliferation of malignant lymphocytes and the development of SLL.

BCL-2: The Critical Axis and Normal Function

While p53 plays a significant role in SLL, its polygenic nature means other genes contribute substantially to the condition. Another gene is *BCL-2* (B-cell leukemia/lymphoma 2 protein) which promotes the survival of cells. *BCL-2* does not increase the rate at which cells divide; rather, it prevents apoptosis. If cell death does not occur, cancer may result. On the other hand, other genes, such as *TP53*, are pro-apoptotic as they promote programmed cell death. This function is activated as a result of harmful events inside or outside the cell, which leads to one of the homologous proteins, BAX (BCL-2–associated X protein), being activated and undergoing structural changes, membrane insertion, and oligomerization. Another protein similar to BAX is BAK (BCL-2 homologues antagonist/killer), which mediates apoptosis. It forms a structure like a channel in the outer membrane of the mitochondria and a mitochondrial protein known as cytochrome c leaves the mitochondria to activate a caspase and apoptosis. However, the BCL-2 proteins also inhibit those that are pro-apoptotic (Klanova and Klener 938). Therefore, cancer progression is determined by the combination of these opposing proteins.

BCL-2: Structure

The structure of BCL-2 is composed of two central hydrophobic helices packed against four amphipathic α-helices. The surface of the protein includes a hydrophobic groove, which is responsible for anti-apoptotic activity. Additionally, BCL-2 is composed of four conserved domains: BH1, BH2, BH3, and BH4. A transmembrane domain is also included which is crucial in regulating apoptosis (Petros et al. 3012 - 7).

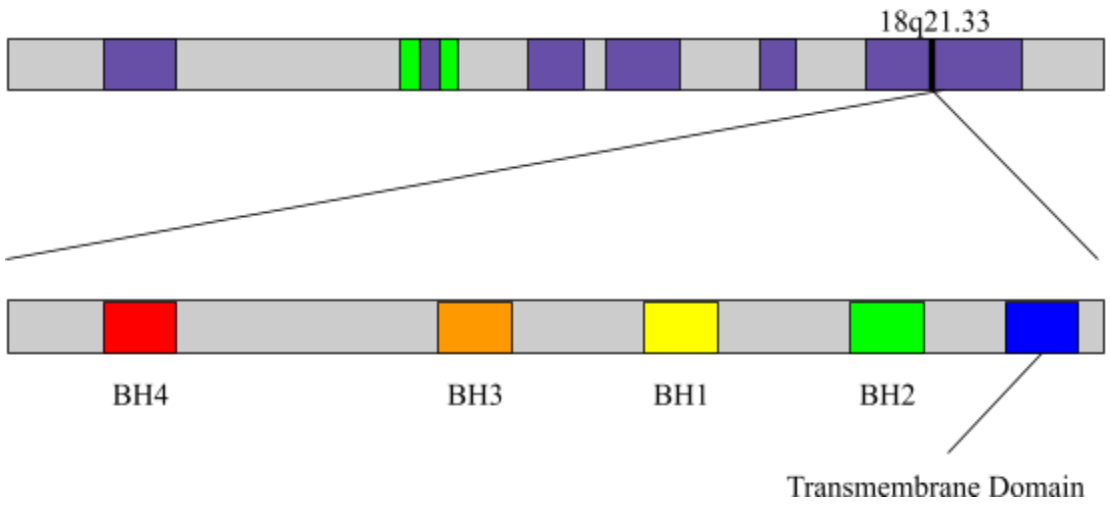


Fig. 2: Domain architecture schematic for BCL-2 (adapted from Karst and Li 318 - 30)

BCL-2: Knockout Mouse

Knockout mice with a *BCL-2* deactivation typically exhibit growth retardation, lymphoid apoptosis, and earlier mortality. This deficiency also disrupted embryogenesis which led to cranial and facial abnormalities in the mice (Veis et al. 229 - 40). Therefore, the deactivation of *BCL-2* is shown to reduce cellular life expectancy due to evidence of disturbed cell growth and physical changes in mice anatomy.

BCL-2: Mutations and their Consequences

The primary mutations in the *BCL-2* gene causing SLL include the frequent overexpression of B-cell leukemia/lymphoma-2 (BCL-2/BCL2), myeloid cell leukemia-1 (MCL-1/MCL1), and B-cell lymphoma-extra large (BCL-XL/BCL2L1). These abnormalities play a vital role in lymphoma pathogenesis, drug resistance, and the progression of disease (Klanova and Klener 938). Different abnormalities exist for different cancer subtypes, in some cases, leading to similar outcomes due to identical issues. Among NHL subtypes, gene translocation and loss of chromosomal regions are common. Overall, these mutations lead to a vast increase in cell life span, with cancer being the consequence.

In studying the oncology behind B Non-Hodgkin Lymphomas (B-NHLs), a type of NHL which begins in B lymphocytes, the presence of the BCL-2 protein is known as BCL-2 positivity. Measuring the extent of this gene expression may be associated with phenotypes such as in SLL. Differing levels of BCL-2 positivity may also be found across subtypes of B-NHL. For example, diffuse large B cell lymphoma (DLBCL) has a BCL-2 positivity around 49% - 67% being overexpressed by *BCL-2* translocation and amplification (Scott et al. 2848 - 56; Hu et al. 4021 - 31; Sehn et al. 131 - 3; Bolen et al. 2298 - 2307). In follicular lymphoma (FL), BCL-2 positivity is typically at least 90%, which is caused by a *BCL-2* translocation (Leich et al. 826 - 34; Bende

et al. 18 - 29; Guo et al. 148 - 52). BCL-2 is positive in mantle cell lymphoma (MCL) caused by the loss of the chromosomal region 13q14.3 or gains in the region 18q21, with the former being more than twice as common as the latter (Guo et al. 148 - 52; Davids et al. 826 - 33; Tagawa et al. 1348 - 58). Similarly, in SLL and CLL, *BCL-2* is highly expressed in most cases due to the loss of the 13q14.3 region, BCL-2 hypomethylation, and translocation between chromosomes 11 and 14 (Majid et al. 874 - 7; Calin et al. 15524 - 9; Hanada et al. 1820 - 8). Translocation of the *BCL-2* gene is a common mechanism underlying the causes of some B-NHL subtypes. In other cases, *BCL-2* may not be as positively expressed. This information is summarized in **Table 3**.

B - NHL subtype	BCL-2 Positivity	BCL-2 Mechanism of Overexpression
DLBCL	49% - 67%	*BCL-2* translocation *BCL-2* amplification
FL	> 90%	*BCL-2* translocation
MCL	BCL-2 positive	Loss of 13q14.3 (55% of cases) Gain of 18q21 (20% of cases)
SLL/CLL	BCL-2 positive (highly expressed in most cases)	Loss of 13q14.3 *BCL-2* hypomethylation *BCL-2* translocation

Table 3: BCL-2 positivity levels and mechanisms of overexpression in B-NHL subtypes

MYC: The Critical Axis and Normal Function

The *MYC* gene is an oncogene and a transcription factor which possesses a variety of functions influencing cellular activities such as the cell cycle, hematopoiesis, and responding to damaged DNA. There are multiple different types of the *MYC* gene, but in this review, we will focus on *c-MYC*. Since *c-MYC* is multifunctional, it is regulated to various extents as well. If *c-MYC* is deregulated, cellular activities such as the cell cycle and DNA repair are not properly regulated, potentially leading to cancer (Ahmadi et al. 121). This is comparable to p53 as when the *TP53* gene is mutated, damaged DNA is not repaired, which may become a cause of cancer due to unfixed genomic alterations. In cell cycle regulation, the *c-MYC* gene plays a role in the cell entering the G1 phase. During G1, the cell grows in size, produces mRNA, and prepares for cell replication. c-MYC is needed in order for the cell to be able to progress into the next stage of cell replication which is the S phase where cellular DNA is replicated. For instance, positive regulators of c-MYC include the kinases ERK, CDK2, and Src (Sears et al. 2501 - 14).

On the other hand, negative regulators include ubiquitin ligases such as Fbw7 (Welcker and Clurman 83 - 93). Collectively, these regulators, along with many others, influence the level of *c-MYC* gene expression during the cell cycle. If the effects of the negative regulators outweigh those of the positive regulators, *c-MYC* is overall deregulated and cancers such as SLL and CLL may result. During the DNA damage response, c-MYC plays a role in the expression of NBS1, a

protein which influences the expression of the protein kinase ATM. Both NBS1 and ATM are involved in the DNA damage response and repair (Guerra et al. 8924). Downstream, ATM also influences p53 which ultimately decides whether the cell is to undergo apoptosis or the DNA will be repaired. Deregulation of *c-MYC* disrupts the coordinated function of ATM and NBS1 in conjunction with p53, thereby impairing the initiation of DNA damage repair or apoptosis in response to genomic damage. Thus, malignant cells may form with mutated DNA resulting in cancer. In initiating apoptosis, c-MYC is involved in the intrinsic and the extrinsic pathways. The intrinsic pathway is initiated by cellular stresses such as DNA damage and hypoxia (Candé et al. 4727 - 34).

In contrast, the extrinsic pathway is activated by the binding of the death ligand to death receptors on the cell surface. Death receptors are a type of tumor necrosis factor receptor family which are responsible for the activation of the signaling cascade and initiation of apoptosis (Lavrik et al. 265 - 7). *c-MYC* deregulation prohibits it from responding to cellular stressors or ligand binding. The signaling cascade will not activate and apoptosis will not be initiated, resulting in cancers such as SLL and CLL. In all, these failures collectively have the same outcome which demonstrates the multiple levels of *c-MYC* gene regulation and their effects.

c-MYC: Structure

The structure of c-MYC consists of an N-terminal transactivation domain (TAD) which includes conserved transcriptional regulation elements known as MYC boxes (MB), a central region, and a C-terminus region with a basic helix–loop–helix leucine zipper domain (bHLHLZ) (Conacci-Sorrell et al. 014357). The transcription and stability of c-MYC depend on the TAD, making it crucial for the gene in its role as a transcription factor. c-MYC transformation depends on the bHLHLZ region, which is why it is the most studied region structurally. This region is also crucial in the design of drugs to inhibit c-MYC (Draeger and Mullen 1785 - 93). The central region contains a protein sequence known as PEST containing the amino acids proline (P), glutamic acid (E), serine (S), and threonine (T) (Kurland and Tansey 3624 - 9). In **Fig. 3**, there are two nuclear localization sequences (NLS) or amino acid sequences that function as indicators for the c-MYC protein transport into the cell nucleus (Rosales et al. 1953).

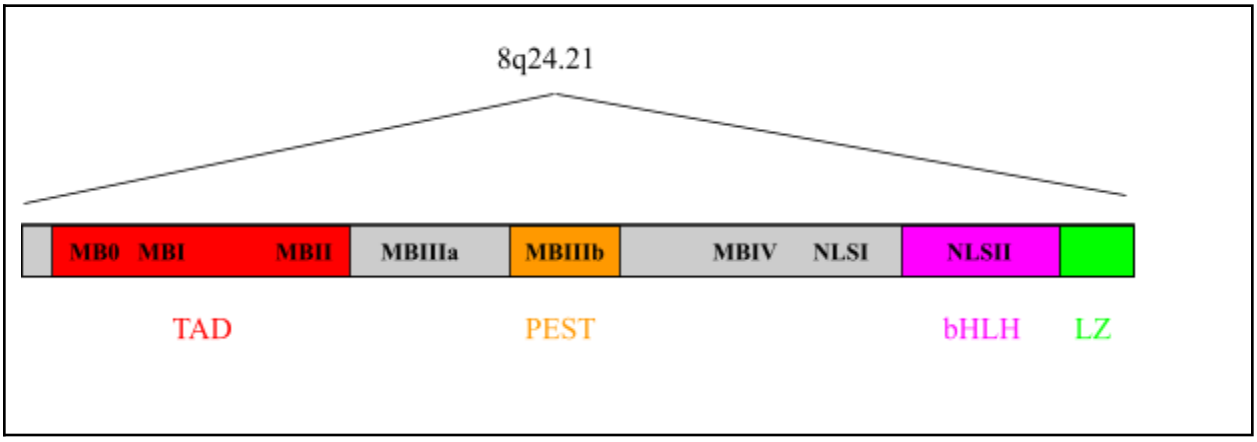


Fig. 3: Domain architecture schematic for c-MYC (adapted from Soucek et al. 1038)
The Soucek group demonstrated that the c-MYC protein is divided into the TAD with c-MYC boxes MB0, MBI, and MBII, the protein sequence PEST with c-MYC box MBIIIb, and the bHLHLZ region with nuclear localization sequence NLSII.

MYC: Knockout Mouse

Knockout mice with haploinsufficiency (a phenomenon when one functional copy of a gene is insufficient to create a normal phenotype) and a deactivated *MYC* gene after birth were found to live considerably longer than the wild-type mice, exhibit premature aging, and experience a reduced incidence of cancer. The deactivation of *MYC* also facilitated a reduction in mitochondrial and ribosomal structural and functional strength, the creation of reactive oxygen species, and cellular deterioration (Prochownik and Wang 1244321). This is unique compared to the findings of deactivating *TP53* and *BCL-2* as those studies found that deactivating those genes decreased the lifespan of the mice. In contrast, deactivating the *MYC* gene benefits the mice more than it harms them. This is due to reduced cell proliferation from the *MYC* deactivation which reduces the incidence of cancer.

c-MYC: Mutations and their Consequences

In a study of 750 patients, a variety of gene variants were identified in association with the diffuse large B cell lymphoma phenotype, including one example, SLL. Most of the gene variants in the population were single-nucleotide substitutions, one example being single-nucleotide variations (SNVs). The SNVs were later classified as either germline single-nucleotide polymorphisms (SNPs) or somatic mutations. In the c-MYC coding sequence, fourteen SNPs were identified, with two of the most prevalent being rs4645959 and rs2070582. When SNPs were excluded, c-MYC mutations were found in 250 patients with DLBCL. These mutations were primarily found in the 5'UTR (untranslated region) and coding sequence regions, whereas mutations in the 3'UTR regions were less frequent. In the two groups in the study, the MUT-MYC (mutant) and the WT-MYC (wildtype), the main clinical difference was that MUT-MYC patients with DLBCL originating within the germinal area of the lymph nodes (GCB DLBCL) had a higher frequency of primary nodal origin (lymph nodes) instead of the lymphatic system (extranodal). On the molecular scale, the mutant group experienced more c-MYC 5'UTR mutations ($P < 0.0001$) and CD10 ($P = 0.0052$) and PI3K expression ($P = 0.048$). Another finding found that the mutant group had more frequent *c-MYC* rearrangements compared to the wildtype group ($P = 0.011$) (Xu-Monette et al. 3593 - 605). These are several examples of the most common mutations in individuals with a mutated *c-MYC* gene in DLBCL cases like SLL.

The p53–BCL-2–c-MYC Connection

The three genes discussed can be integrated into a unified model that represents the progression of cancer at a genetic level. First, when *c-MYC* is deregulated, p53-dependent apoptosis is promoted as well as the p53-independent activation of the BH3-only protein Bim

(pro-apoptotic). Second, to promote tumor development, targeted inactivation of the p53 or the Bim pathway in B cells that express c-MYC is sufficient (Egle et al. 6164 - 9; Schmitt et al. 2670 - 7). Likewise, tumor-derived c-MYC mutants that exhibit weakened Bim activation but normal increases in p53 levels are capable of creating tumors (Hemann et al. 807 - 11). Therefore, p53-independent and dependent signals act together to promote apoptosis. These signals combined must outweigh the opposing antiapoptotic BCL-2 group, as the inactivation of one of the pro-apoptotic effectors may drop the system below the firing threshold required for apoptosis which may lead to cancer such as SLL and CLL (Hemann and Lowe 1256 - 9).
Note: Puma, Noxa, and Bax are pro-apoptotic proteins known for their roles in the cell signaling pathway that regulates apoptosis. This information is visualized in **Fig. 4**.

The Pathology and Cytokines of SLL and CLL

Pathology is the study of disease which concerns its causes and its effects. A pathologist establishes a diagnosis of SLL or CLL through examination of a tissue biopsy or blood stain.

To begin, we will compare the appearance of normal and cancerous lymph node tissue (SLL) to understand the differences in the tissue specifically that warrant an SLL diagnosis. In a normal lymph node tissue under a microscope, several components must be distinctly visible including the capsule (connective tissue surrounding the node), cortex (outer region of the node), and medulla (internal part of the node). Furthermore, in a normal lymph node, (Fig. 5) there can be expected a variety of cells such as lymphocytes and plasma cells. On the other hand, a tissue with SLL would exhibit the destruction of normal lymphoid tissue by small lymphocytes which contain irregularly condensed chromatin and round and relatively similar-looking nuclei (Fenu and Rosenthal 32010762). Additionally, paler centers of proliferation are scattered in different areas in the tissue and the cells are characterized as being a more uniform population of B lymphocytes. These differences are visually demonstrated in **Fig. 5** (retrieved from Sangle and Pernick (a) and Kajtar (b)).

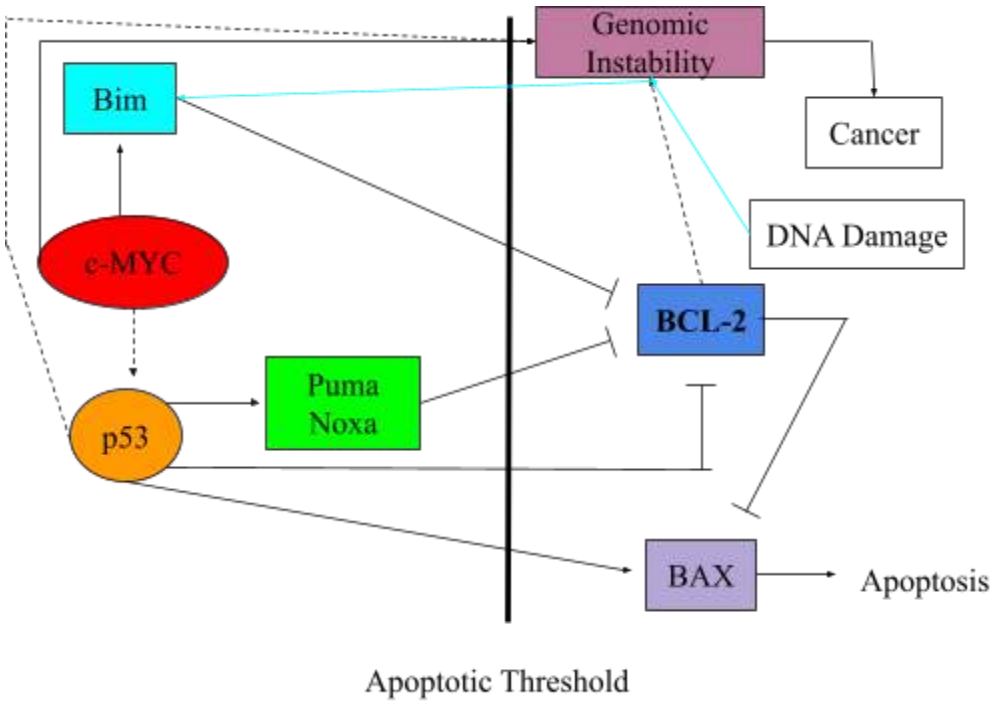
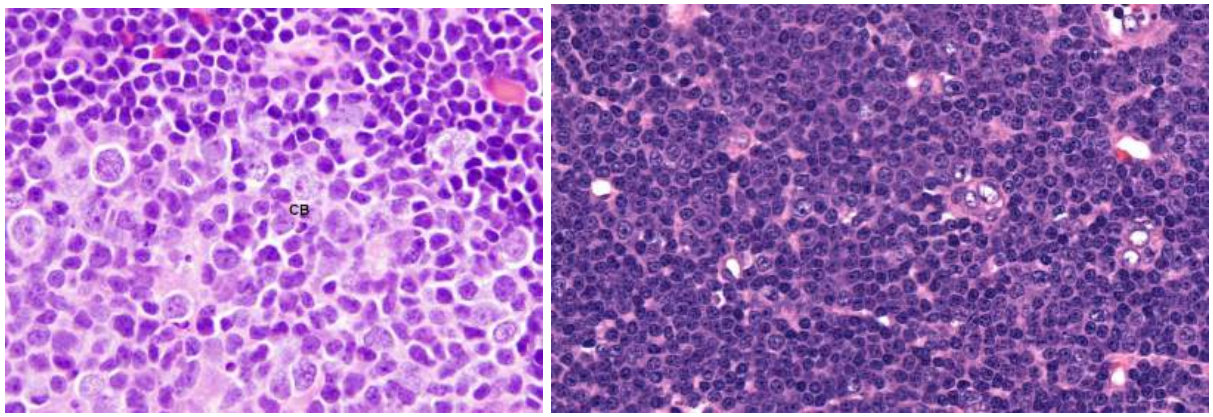


Fig. 4: The c-MYC, p53, and BCL-2 pathway in apoptosis (adapted from Hemann and Lowe 1256 - 9)

Hemann and Lowe identified the proteins that induce or inhibit one another ultimately leading to cancer or apoptosis.



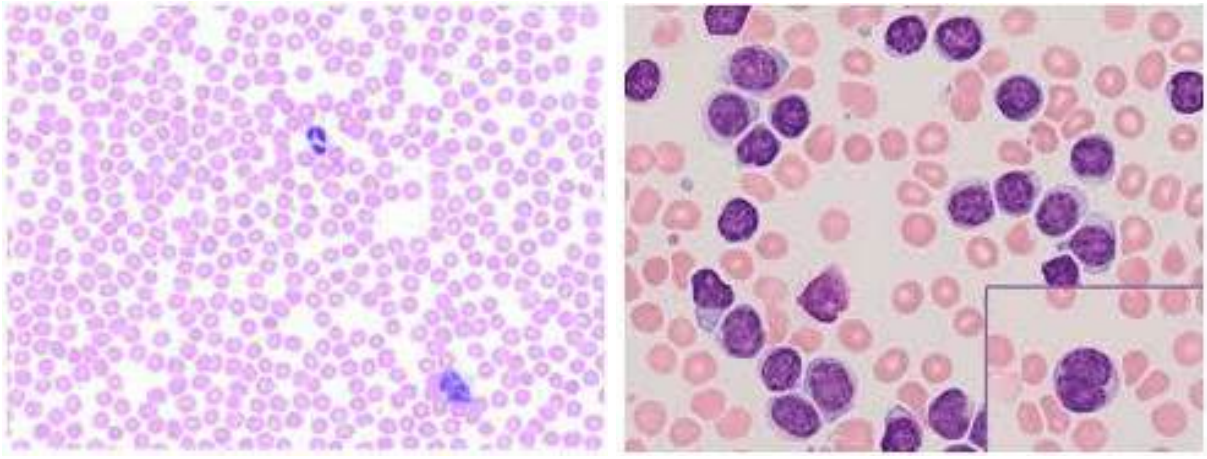
a b

Fig. 5: Histological slide images of lymph node tissues

a. Normal (retrieved from Sangle and Pernick)
b. Cancerous; SLL (retrieved from Kajtar)

When considering the histology of CLL, a sample of the patient's blood must be analyzed as it is a blood disorder and does not concern tissues. A normal blood smear would be

characterized as having relatively uniform RBCs in size and shape as well as a pale center. WBCs should also be present in normal amounts and appear mature without abnormalities. Last, there should not be any smudge cells present. In contrast, a peripheral blood smear of a patient with CLL would include small and mature lymphocytes with darkly stained nuclei, but in greater numbers due to the presence of cancer. Condensed chromatin and smudge cells, which are more fragile than normal lymphocytes, should also be present (Hallek et al. 2745 - 60; Nowakowski et al. 449 - 53). These details are illustrated in the images shown in **Fig. 6** (retrieved from Kachur and Rosen (a) and McCall et al. (b)).



a b

Fig. 6: Hematological slide images of peripheral blood smears
a. Normal (retrieved from Kachur and Rosen 33889717)
b. Cancerous; CLL (retrieved from McCall et al. 9185432)

Cytokines are signaling proteins involved in the process of cell-to-cell communication, and are especially vital in the immune system. Cytokines are either up-or downregulated in SLL and CLL. Some examples of overexpressed cytokines compared to the healthy control group without SLL include: β2M, I-309, IL-5, IL-10, IL-16, IL-28A, and MCP-1. In contrast, instances of underexpressed cytokines compared to the control include the following: Fractalkine, IFN-γ, IL-17, IL-23, IL-33, TGF-α, and VEGF (Karmali et al. 161 - 70). Despite the expression patterns being mixed, an overall increase in Th2 cytokines and a decrease in Th1 and Th17 cytokines is observed. This abnormal combination of cytokine levels mediates the inflammatory response, which facilitates cancer development. The data is summarized in **Table 4**.

The Influence of Drugs on SLL and CLL Recovery

A variety of drugs and procedures are being developed to treat patients with SLL and CLL. Combinations of drugs and other treatments may also be utilized in other NHL subtypes. The data presented in the following paragraphs is summarized in **Table 5**.

For example, R-CHOP stands for rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone, which is a treatment regimen that uses the five mentioned drugs to treat different

Overexpressed cytokines	P for means	Underexpressed cytokines	P for means
β2M	< 0.001	Fractalkine	= 0.006
I-309	= 0.008	IFN-γ	= 0.003
IL-5	= 0.001	IL-17	= 0.009
IL-10	= 0.001	IL-23	= 0.02
IL-16	< 0.001	IL-33	= 0.008
IL-28A	= 0.001	TGF-α	< 0.001
MCP-1	< 0.001	VEGF	= 0.008

Table 4: Over and Underexpressed cytokines in SLL

NHL subtypes. One study found that four cycles of R-CHOP with two additional doses of rituximab had a 94% rate of progression-free survival (PFS) of three years. This was compared to six cycles of R-CHOP, which had a 96% rate of PFS of three years. Cycles were repeated every 21 days with a total sample size of 588 patients. These findings were not statistically significant however so the difference observed between these two treatment options is likely attributable to chance (Poeschel et al. 2271 - 81). Due to the heterogeneous responses to the drugs, further interrogation of these trials is warranted based on the results.

Another study utilized DA-EPOCH-R, which stands for dose-adjusted etoposide, vincristine, doxorubicin with cyclophosphamide, prednisone, and rituximab, which is another treatment regimen used to treat NHL. Six cycles of DA-EPOCH-R were found to have a two-year PFS rate of 79%. This was compared to six cycles of R-CHOP, which was found to have a 76% rate of two-year PFS (P = 0.65). All cycles were repeated every 21 days with 491 patients. Like the previous study, these results were not statistically significant either (Bartlett et al. 1790 - 9). This indicates little to no statistical difference between the rates of PFS of the two treatment methods, which warrants further work on these treatment techniques.

A further example of a drug used to treat cancer is Obinutuzumab, which is a prescription medicine and also a monoclonal antibody. When studied in combination, the Obinutuzumab and 6 or 8 cycles of CHOP produced a three-year PFS rate of 70%. This was compared to a 67% rate of three-year PFS with 6 or 8 cycles of just R-CHOP (P = 0.39). The cycles were repeated every 21 days across a sample of 1,418 patients. Ultimately, this difference was also not statistically significant (Vitolo et al. 3529 - 37). Therefore, there exists relative similarity between these treatment methods and further research should be done to determine a more effective treatment.

In addition to the above discussed combination drug therapies, Venetoclax is a biological small molecular inhibitor designed to target BCL-2. Mechanistically, Venetoclax targets the BH-3 domain of BCL-2 (Perini et al. 66; Itchaki and Brown 270 - 87). Venetoclax functions in a similar manner to BH-3 which fits perfectly into the space of the BCL-2 molecule where BH-3 binds. Proteins known as BIM actively promote apoptosis which attach to the BCL-2 protein. Consequently, the presence of proteins known as BH-3 only proteins also initiate apoptosis. The proteins can activate other proteins such as BAX and BAK which are responsible for inducing apoptosis. BH-3 only proteins can also prohibit proteins that attempt to prevent apoptosis, like MCL-1. Therefore, Venetoclax can release proteins that trigger apoptosis and prevent tumor development (Mihalyova et al. 10 - 25; Smith et al. 2150 - 63; Merino et al. 879 - 91). A study's findings also demonstrated that in cells with T-cell acute lymphoblastic leukemia (T-ALL), cells treated with Venetoclax showed a considerable increase in apoptosis compared to the control group ($P < 0.05$). The same was observed in cells with B-cell acute lymphoblastic leukemia (B-ALL) treated with Venetoclax ($P < 0.05$) (Tari et al. 229 - 41).

Treatment comparison	Rate of 2- or 3-year PFS	Statistical Significance
R-CHOP x 4 (+2 additional rituximab) R-CHOP x 6	3 year PFS 94% 3 year PFS 96%	No - neither treatment is inferior to the other
DA-EPOCH-R x 6 R-CHOP x 6	2 year PFS 79% 2 year PFS 76%	No - neither treatment is inferior to the other
Obinutuzumab + CHOP x 6 - 8 R-CHOP x 6 - 8	3 year PFS 70% 3 year PFS 67%	No - neither treatment is inferior to the other

Table 5: Comparing rates of PFS among different treatment regimens

The Influence of Immunotherapies on SLL and CLL Recovery

In addition to the wide range of prescription medicines available for SLL and CLL treatment, immunotherapies are also available to treat these medical conditions should the previous options be insufficient to resolve the issue. Some examples are chimeric antigen receptor (CAR) T-cell therapy and immune checkpoint inhibitors.

CAR T-cell therapy involves collecting T lymphocytes from the patient, and genetically altering them to express a chimeric antigen receptor (referred to as CAR) that targets a specific cancer protein on cancerous cells. Then the modified cells are returned into the patient and used to fight the cancer. The anatomy of the CAR contains an antigen recognition domain targeting a specific antigen, a hinge connecting the recognition site to the transmembrane domain, and an intracellular domain with a CD3z chain vital for the signaling of the T cell receptor (Brentjens et al. 279 - 86; Geldres et al. 3 - 9; Brudno and Kochenderfer 31 - 46; Kochenderfer et al. 689 - 702; Sadelain 3392 - 400). One target for treating lymphomas is the type 1 transmembrane glycoprotein CD19. It is expressed on the surface of B lymphocytes as well as on newly

transformed cancer cells, like in B-NHL, in more than 95% of cases (Ramos et al. 112 - 8; Bailly et al. 505 - 17). Thus, CD19 is a viable target for treating these lymphomas.

Despite the potential effectiveness of CAR T-cell therapy, several risks are associated with this technique. For example, cytokine release syndrome (CRS), a systemic inflammatory response, is the most frequent adverse effect. The exact mechanism by which this occurs is not entirely understood, although it is related to the release of inflammatory cytokines (Gauthier and Turtle 50 - 2). The main symptoms include fever, fatigue, rigors, or anorexia, which may be followed by tachycardia, hypoxia, neurological changes, and more (Riches and Gribben 207 - 35). Another obstacle is immune effector cell-associated neurotoxicity syndrome (ICANS), a side effect characterized by inflammation found in the brain. Again, the mechanisms by which this occurs are not entirely understood, but some theories claim it may be associated with CNS T-cell trafficking, greater vascular permeability, and a disruption in the blood-brain barrier (Lee et al. 517 - 28). Symptoms of this include headaches, focal deficits, confusion, seizures, and hallucinations (Gust et al. 1404 - 19; Wang et al. 896 - 925). A third challenge is infection. The therapy weakens the patient's immune system, resulting in a depletion of B lymphocytes, which can lead to diseases such as adenovirus, cytomegalovirus (CMV), Epstein–Barr virus (EBV), and human herpesvirus 6 (HHV6) (Ostojska et al. 1595; Liu et al. 126 - 9).

In contrast to CAR T-cell therapy, immune checkpoint inhibitors block proteins known as checkpoints that regulate the immune response to cancer cells. Cancer cells may exploit these checkpoints to avoid destruction. One example of an immune checkpoint inhibitor is Bruton's Tyrosine Kinase (BTK). BTK has been identified as a part of the B-cell receptor (BCR) signaling pathway and is necessary for normal B-cell development (Singh et al. 57; Hendriks et al. 219 - 32). Some examples of BTK inhibitors specifically for the treatment of SLL and CLL include ibrutinib and acalabrutinib. These inhibitors deactivate BTK by binding to Cys481 on the ATP-binding region of BTK, which inactivates the BTK enzyme and disrupts BCR signaling. This signaling disruption can then prevent the disease progression (Singh et al. 57; Palma et al. 686768).

There are a variety of adverse effects, many of which are shared among three immune checkpoint inhibitors: ibrutinib, acalabrutinib, and zanubrutinib. The most common effects include diarrhea, arthralgia, hypertension, atrial fibrillation, UTI, dyspepsia, headache, cough, and bleeding (Byrd et al. 3441 - 52). **Table 6** compares CAR T-cell therapy to immune checkpoint inhibitors based on their operational mechanisms and adverse effects.

Cancer Diagnostics

Even though research efforts towards developing drugs are taking shape and are successful to an extent, diagnosing a patient with cancer as accurately and early as possible gives the best chance to healthcare professionals in treating the disease effectively.

One group of tests that can be done to diagnose a patient early on with CLL is a blood test. Notable in this group are the complete blood count (CBC) and the peripheral blood smear.

Type of therapy	CAR T-cell therapy	Immune checkpoint inhibitors
Operational mechanism	Genetically altering T-lymphocytes	Blocking checkpoints that regulate immune response
Risks/adverse effects	Fever, fatigue, anorexia, tachycardia, hypoxia, neurological changes (eg. cytokine release syndrome which is referred to as CRS), headaches, seizures, hallucinations, neurological toxicity (eg. ICANS), adenovirus, CMV, EBV, HHV6	Diarrhea, nausea, bleeding, fatigue, cough, neutropenia, hypertension, atrial fibrillation, headache, muscle spasms, urinary tract infection

Table 6: The operational mechanisms and risks of CAR T-cell therapy and immune checkpoint inhibitors

First, the CBC test measures the amount of WBCs, RBCs, and platelets in the bloodstream. This is also done with a differential which counts the types of WBCs in the sample. When a blood-related issue is suspected, these tests are among the first to be done. Individuals with CLL exhibit lymphocytosis, characterized by an excessive number of lymphocytes. However, other tests should be done for improved accuracy as CLL patients may also display low levels of RBCs and platelets. Second, the peripheral blood smear consists of a microscopic analysis of the blood sample. In CLL, the blood smear may show numerous smudge cells, or abnormally shaped or damaged lymphocytes (Hallek et al. 2745 - 60; Nowakowski et al. 449 - 53).

Biopsies are another method to detect an early onset of cancer. In a biopsy procedure, a sample of tissue is removed for microscopic analysis to diagnose or monitor a disease. For example, a lymph node biopsy involves removing a portion of the lymph node to be tested for leukemia cells. This method is typically utilized when blood tests are insufficient to make the diagnosis or when the doctor wants to know if the leukemia has progressed into a more aggressive cancer. If these methods are combined and still insufficient for diagnosis, a bone marrow aspiration and biopsy are performed together. In the bone marrow aspiration, a hollow needle is inserted into the bone marrow of the hip bone, and a syringe suctions a small quantity of liquid bone marrow. The biopsy is performed by taking a small piece of bone and marrow using a larger needle that is inserted into the bone (Hallek et al. 2745 - 60).

These blood tests and biopsies are viewed microscopically by a pathologist, hematologist, or oncologist. Pathologists specialize in examining specimens to study diseases. Hematologists specialize in the detection and treatment of blood-related disorders. Oncologists specialize in the detection and treatment of cancer.

Flow cytometry and immunohistochemistry are two other tests used to classify cancer cells based on the antigens (tags) on their surfaces. This is known as immunophenotyping. Both tests involve treating samples of cells with antibodies, which are proteins that bind to specific antigens. Flow cytometry categorizes the treated cells using a machine that analyzes their physical and chemical features with the aid of a laser (Hallek et al. 2745 - 60). On the other hand,

in immunohistochemistry, treated cells are viewed under a microscope to determine if antibodies are bound to antigens on the cell surfaces (DiGiuseppe and Borowitz 6 - 10).

Chromosome, molecular, and genetic tests are also used to understand the causes of cancer phenotypes. For example, in cytogenetics, cells are grown in a laboratory for several days before their chromosomes are analyzed microscopically. Another assay is fluorescent *in situ* hybridization, which uses fluorescent dyes that bind to regions of specific chromosomes. This technique is typically faster than cytogenetics because cells do not need to be grown. These tests usually detect alterations in an individual's genome. For instance, a section of a chromosome may be missing (deletion), an extra copy of a chromosome may be present (trisomy), or genetic material is swapped between two chromosomes (translocation). Some examples of genomic alterations in CLL include del(11p) or del(17p), which are deletions (Molica et al. 1177 - 85). Regarding molecular tests, common proteins analyzed include immunoglobulins. These antibodies aid the body in fighting infections and are composed of light and heavy chains. The immunoglobulin heavy chain variable region genes, IGHV or IgVH, and whether they are mutated, can provide insight into the aggressiveness of CLL (Thompson et al. 303 - 9). As previously mentioned, mutations in the *TP53* gene may also provide insight into the severity of not only CLL, but also other cancers.

Some blood-based tests that can be applied with the help of biomarkers to detect cancer are circulating tumor DNA (ctDNA) and circulating tumor cells (CTC). ctDNA is composed of DNA derived from tumor and normal cells in extracellular bodily fluids. The biomarker can be used to assess the molecular basis of the tumor and predict the patient's future cancer progression (Mithraprabhu and Spencer 129 - 45). In comparison, CTCs are released from tumors and are moved through the body via the circulatory system or lymphatic vessels. Afterwards, they may settle and form micrometastases (small, undetectable clusters of cancer cells) under the right conditions. This information may also describe the tumor's molecular landscape, assisting in cancer diagnosis (Vidlarova et al. 3902).

Conclusion: Future Perspectives

The variability in the statistical significance of treatment outcomes across different therapeutic agents for NHL underscores the uneven progress in drug development within this field. While some therapies demonstrate clear clinical benefit, others yield more modest or uncertain improvements. However, this must be understood in the context that immense time and effort are invested in developing new treatments for these cancers, and not all of them will yield a notable improvement compared to other similar drugs or courses of medication. Importantly, identifying and addressing limitations in new treatments remains a critical component of the drug development process. Lessons learned from suboptimal outcomes are invaluable in guiding future research directions and refining treatment strategies to enhance efficacy and patient care.

Researchers funded by the National Cancer Institute are currently advancing knowledge about lymphomas. Some current treatment techniques include chemotherapy, radiation therapy, targeted therapy, and immunotherapy to treat all kinds of NHL cancers, including SLL and CLL.

Another treatment influences the signaling pathways, such as the B-cell receptor pathway, which regulates NHL. To treat SLL, a new drug, ibrutinib (Imbruvica), has been approved to shut down the pathway. As technology advances, we will undoubtedly see more advanced and effective methods of treating various cancers. Despite this, it is still paramount to maintain a healthy lifestyle as cancer prevention is more beneficial for any individual, regardless of whether or not they have cancer. Ultimately, the combination of proactive health decisions and a detailed understanding of biological science seems to be the foundation for effective cancer management.

Works Cited

Ahmadi, Seyed Esmaeil, et al. "Myc: A Multipurpose Oncogene with Prognostic and Therapeutic Implications in Blood Malignancies." Journal of Hematology & Oncology, U.S. National Library of Medicine, 9 Aug. 2021, pmc.ncbi.nlm.nih.gov/articles/PMC8351444/#Sec5.

Babamohamadi, Mehregan, et al. "Recent findings on the role of wild-type and mutant p53 in cancer development and therapy." Frontiers in Molecular Biosciences, U.S. National Library of Medicine, 26 Sept. 2022, https://pubmed.ncbi.nlm.nih.gov/36225257/.

Bailly, Sarah, et al. "Targeting CD19 in Diffuse Large B-Cell Lymphoma: An Expert Opinion Paper." Hematological Oncology, U.S. National Library of Medicine, 7 June 2022, pubmed.ncbi.nlm.nih.gov/35488888/.

Bartlett, Nancy L, et al. "Dose-Adjusted Epoch-R Compared with R-CHOP as Frontline Therapy for Diffuse Large B-Cell Lymphoma: Clinical Outcomes of the Phase III Intergroup Trial Alliance/CALGB 50303." Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, U.S. National Library of Medicine, 20 July 2019, pmc.ncbi.nlm.nih.gov/articles/PMC6774813/.

Beaulieu, Marie-Eve, et al. "Structural and Biophysical Insights into the Function of the Intrinsically Disordered Myc Oncoprotein." Cells, U.S. National Library of Medicine, 22 Apr. 2020, pmc.ncbi.nlm.nih.gov/articles/PMC7226237/.

Bende, RJ, et al. "Molecular Pathways in Follicular Lymphoma." Leukemia, U.S. National Library of Medicine, 12 Oct. 2006, pubmed.ncbi.nlm.nih.gov/17039231/.

Bolen, CR, et al. "Prognostic Impact of Somatic Mutations in Diffuse Large B-Cell Lymphoma and Relationship to Cell-of-Origin: Data from the Phase III Goya Study." Haematologica, U.S. National Library of Medicine, 1 Sept. 2020, pubmed.ncbi.nlm.nih.gov/33054054/.

Brentjens, Renier J, et al. "Eradication of Systemic B-Cell Tumors by Genetically Targeted Human T Lymphocytes Co-Stimulated by CD80 and Interleukin-15." Nature Medicine, U.S. National Library of Medicine, 10 Feb. 2003, pubmed.ncbi.nlm.nih.gov/12579196/.

Brudno, Jennifer N, and James N Kochenderfer. "Chimeric Antigen Receptor T-Cell Therapies for Lymphoma." Nature Reviews. Clinical Oncology, U.S. National Library of Medicine, 31 Aug. 2017, pubmed.ncbi.nlm.nih.gov/28857075/.

Byrd, John C, et al. "Acalabrutinib versus Ibrutinib in Previously Treated Chronic Lymphocytic Leukemia: Results of the First Randomized Phase III Trial." Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, U.S. National Library of Medicine, 26 July 2021, pubmed.ncbi.nlm.nih.gov/34310172/.

Calin, GA, et al. "Frequent Deletions and Down-Regulation of Micro- RNA Genes Mir15 and MIR16 at 13q14 in Chronic Lymphocytic Leukemia." Proceedings of the National Academy of Sciences of the United States of America, U.S. National Library of Medicine, 14 Nov. 2002, pubmed.ncbi.nlm.nih.gov/12434020/.

Candé, Celine, et al. “Apoptosis-Inducing Factor (AIF): Key to the Conserved Caspase-Independent Pathways of Cell Death?” Journal of Cell Science, U.S. National Library of Medicine, 15 Dec. 2002, pubmed.ncbi.nlm.nih.gov/12432061/.

Conacci-Sorrell, Maralice, et al. “An overview of MYC and its interactome.” Cold Spring Harbor Perspectives in Medicine, U.S. National Library of Medicine, 1 Jan. 2014, https://pubmed.ncbi.nlm.nih.gov/24384812/.

Davids, MS, et al. “Phase I First-in-Human Study of Venetoclax in Patients with Relapsed or Refractory Non-Hodgkin Lymphoma.” Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, U.S. National Library of Medicine, 17 Jan. 2017, pubmed.ncbi.nlm.nih.gov/28095146/.

DiGiuseppe, Joseph A, and Michael J Borowitz. “Clinical Utility of Flow Cytometry in the Chronic Lymphoid Leukemias.” Seminars in Oncology, U.S. National Library of Medicine, 25 Feb. 1998, pubmed.ncbi.nlm.nih.gov/9482521/.

Draeger, LJ, and G P Mullen. “Interaction of the bHLH-ZIP Domain of c-Myc with H1-Type Peptides. Characterization of Helicity in the H1 Peptides by NMR.” The Journal of Biological Chemistry, U.S. National Library of Medicine, 21 Jan. 1994, pubmed.ncbi.nlm.nih.gov/8294427/.

Egle, Alexander, et al. “Bim Is a Suppressor of Myc-Induced Mouse B Cell Leukemia.” Proceedings of the National Academy of Sciences of the United States of America, U.S. National Library of Medicine, 12 Apr. 2004, pubmed.ncbi.nlm.nih.gov/15079075/.

Fenu, Elena M., and Nancy S. Rosenthal. “Educational Case: Small Lymphocytic Lymphoma: Diagnostic Features and Prognosis.” Academic Pathology, U.S. National Library of Medicine, 21 Jan. 2020, https://pmc.ncbi.nlm.nih.gov/articles/PMC6974742/.

Gauthier, Jordan, and Cameron J Turtle. “Insights into Cytokine Release Syndrome and Neurotoxicity after CD19-Specific Car-T Cell Therapy.” Current Research in Translational Medicine, U.S. National Library of Medicine, 3 Apr. 2018, pubmed.ncbi.nlm.nih.gov/29625831/.

Geldres, Claudia, et al. “Chimeric Antigen Receptor-Redirected T Cells Return to the Bench.” Seminars in Immunology, U.S. National Library of Medicine, 12 Jan. 2016, pubmed.ncbi.nlm.nih.gov/26797495/.

Guerra, Lina, et al. “Myc is required for activation of the ATM-dependent checkpoints in response to DNA damage.” Public Library of Science One, U.S. National Library of Medicine, 27 Jan. 2010, https://pubmed.ncbi.nlm.nih.gov/20111719/.

Guo, Y, et al. “BCL2-Negative Follicular Lymphomas Frequently Have BCL6 Translocation and/or BCL6 or P53 Expression.” Pathology International, U.S. National Library of Medicine, 5 Mar. 2007, pubmed.ncbi.nlm.nih.gov/17295647/.

Gust, Juliane, et al. “Endothelial Activation and Blood-Brain Barrier Disruption in Neurotoxicity after Adoptive Immunotherapy with CD19 Car-T Cells.” Cancer Discovery, U.S. National Library of Medicine, 12 Oct. 2017, pubmed.ncbi.nlm.nih.gov/29025771/.

Hallek, Michael, et al. “IWCLL Guidelines for Diagnosis, Indications for Treatment, Response Assessment, and Supportive Management of Cll.” Blood, U.S. National Library of Medicine, 14 Mar. 2018, pubmed.ncbi.nlm.nih.gov/29540348/.

Hanada, M, et al. “Bcl-2 Gene Hypomethylation and High-Level Expression in B-Cell Chronic Lymphocytic Leukemia.” Blood, U.S. National Library of Medicine, 15 Sept. 1993, pubmed.ncbi.nlm.nih.gov/8104532/.

Hemann, M T, and S W Lowe. “The P53-Bcl-2 Connection.” Cell Death and Differentiation, U.S. National Library of Medicine, 1 Oct. 2015, pmc.ncbi.nlm.nih.gov/articles/PMC4590992/#:~:text=p53%20may%20also%20directly%20impact,permeabilization%20of%20mitochondria%20and%20apoptosis.&text=Consequently%2C%20p53%2C%20itself%2C%20may,p53%E2%80%93Bcl%2D2%20interaction.&text=Thus%2C%20tumor%2Dderived%20p53%20mutants,induce%20apoptosis%20at%20the%20mitochondria.&text=Thus%2C%20p53%20can%20regulate%20Bcl,engage%20the%20cell%20death%20machinery

Hemann, Michael T, et al. “Evasion of the P53 Tumour Surveillance Network by Tumour-Derived Myc Mutants.” Nature, U.S. National Library of Medicine, 11 Aug. 2005, pubmed.ncbi.nlm.nih.gov/16094360/.

Hendriks, Rudi W, et al. “Targeting Bruton’s Tyrosine Kinase in B Cell Malignancies.” Nature Reviews. Cancer, U.S. National Library of Medicine, 14 Apr. 2014, pubmed.ncbi.nlm.nih.gov/24658273/.

Hu, S, et al. “Myc/BCL2 Protein Coexpression Contributes to the Inferior Survival of Activated B-Cell Subtype of Diffuse Large B-Cell Lymphoma and Demonstrates High-Risk Gene Expression Signatures: A Report from the International DLBCL RITUXIMAB-CHOP Consortium Program.” Blood, U.S. National Library of Medicine, 28 Feb. 2013, pubmed.ncbi.nlm.nih.gov/23449635/.

Itchaki, Gilad, and Jennifer R Brown. “The Potential of Venetoclax (ABT-199) in Chronic Lymphocytic Leukemia.” Therapeutic Advances in Hematology, U.S. National Library of Medicine, 8 July 2016, pubmed.ncbi.nlm.nih.gov/27695617/.

Kachur, Megan E., and Benjamin J. Rosen. “Educational Case: Anemia in a Neonate.” Academic Pathology, U.S. National Library of Medicine, 8 Apr. 2021, https://pubmed.ncbi.nlm.nih.gov/33889717/.

Kajtár B. “Lymphoma & related disorders Mature B cell neoplasms Small B cell lymphomas with a circulating component CLL / SLL.” PathologyOutlines.com website, 24 Apr. 2024, https://www.pathologyoutlines.com/topic/lymphomaCLL.html.

Karmali, Reem, et al. “Aggressive Disease Defined by Cytogenetics Is Associated with Cytokine Dysregulation in CLL/SLL Patients.” Journal of Leukocyte Biology, U.S. National Library of Medicine, Jan. 2013, pmc.ncbi.nlm.nih.gov/articles/PMC3525834/.

Karst, A M, and G Li. “Bh3-Only Proteins in Tumorigenesis and Malignant Melanoma.” Cellular and Molecular Life Sciences : CMLS, U.S. National Library of Medicine, Feb. 2007, pubmed.ncbi.nlm.nih.gov/17192811/.

Kastenhuber, Edward R, and Scott W Lowe. “Putting P53 in Context.” Cell, U.S. National Library of Medicine, 7 Sept. 2017, pmc.ncbi.nlm.nih.gov/articles/PMC5743327/.
KH, Kubbutat, et al. “Regulation of P53 Stability by MDM2.” Nature, U.S. National Library of Medicine, 15 May 1997, pubmed.ncbi.nlm.nih.gov/9153396/.
Klanova, Magdalena, and Pavel Klener. “Bcl-2 Proteins in Pathogenesis and Therapy of B-Cell Non-Hodgkin Lymphomas.” Cancers, U.S. National Library of Medicine, 10 Apr. 2020, pmc.ncbi.nlm.nih.gov/articles/PMC7226356/.
Kochenderfer, James N, et al. “Construction and Preclinical Evaluation of an Anti-CD19 Chimeric Antigen Receptor.” Journal of Immunotherapy (Hagerstown, Md. : 1997), U.S. National Library of Medicine, Sept. 2009, pubmed.ncbi.nlm.nih.gov/19561539/.
Kurland, John F., and William P. Tansey. “Myc-Mediated Transcriptional Repression by Recruitment of Histone Deacetylase.” Cancer Research, U.S. National Library of Medicine, 15 May 2008, pubmed.ncbi.nlm.nih.gov/18483244/.
LA, Donehower. “The P53-Deficient Mouse: A Model for Basic and Applied Cancer Studies.” Seminars in Cancer Biology, U.S. National Library of Medicine, 7 Oct. 1996, pubmed.ncbi.nlm.nih.gov/9110404/.
Lavrik, Inna, et al. “Death Receptor Signaling.” Journal of Cell Science, U.S. National Library of Medicine, 15 Jan. 2005, pubmed.ncbi.nlm.nih.gov/15654015/.
Lee, Daniel W, et al. “T Cells Expressing CD19 Chimeric Antigen Receptors for Acute Lymphoblastic Leukaemia in Children and Young Adults: A Phase 1 Dose-Escalation Trial.” Lancet (London, England), U.S. National Library of Medicine, 13 Oct. 2014, pubmed.ncbi.nlm.nih.gov/25319501/.
Leich, E, et al. “Follicular Lymphomas with and without Translocation T(14;18) Differ in Gene Expression Profiles and Genetic Alterations.” Blood, U.S. National Library of Medicine, 26 May 2009, pubmed.ncbi.nlm.nih.gov/19471018/.
Lindahl, Lise M, et al. “Stat5 Induces Mir-21 Expression in Cutaneous T Cell Lymphoma.” Oncotarget, U.S. National Library of Medicine, 19 July 2016, pmc.ncbi.nlm.nih.gov/articles/PMC5216756/.
Liu, Wei, et al. “Risk of Hepatitis B Reactivation Is Controllable in Patients with B-Cell Lymphoma Receiving Anti-Cd19 Car T Cell Therapy.” British Journal of Haematology, U.S. National Library of Medicine, 15 July 2020, pubmed.ncbi.nlm.nih.gov/32671820/.
Majid, A, et al. “BCL2 Expression in Chronic Lymphocytic Leukemia: Lack of Association with the BCL2 938A>C Promoter Single Nucleotide Polymorphism.” Blood, U.S. National Library of Medicine, 24 Oct. 2007, pubmed.ncbi.nlm.nih.gov/17959858/.
McCall, Robert K., et al. “Non-Nodal CD5-Negative Mantle Cell Lymphoma with Secondary TP53 Deletion.” Case Reports in Hematology, U.S. National Library of Medicine, 18 Mar. 2020, https://pubmed.ncbi.nlm.nih.gov/32257467/.
Merino, Delphine, et al. “Bh3-Mimetic Drugs: Blazing the Trail for New Cancer Medicines.” Cancer Cell, U.S. National Library of Medicine, 10 Dec. 2018, pubmed.ncbi.nlm.nih.gov/30537511/.

Mihalyova, Jana, et al. “Venetoclax: A New Wave in Hematooncology.” Experimental Hematology, U.S. National Library of Medicine, 2 Mar. 2018, pubmed.ncbi.nlm.nih.gov/29477371/.

Mithraprabhu, Sridurga, and Andrew Spencer. “Analysis of Circulating Tumor DNA.” Methods in Molecular Biology (Clifton, N.J.), U.S. National Library of Medicine, 2018, pubmed.ncbi.nlm.nih.gov/29797256/.

Molica, Stefano, et al. “External Validation on a Prospective Basis of a Nomogram for Predicting the Time to First Treatment in Patients with Chronic Lymphocytic Leukemia.” Cancer, U.S. National Library of Medicine, 7 Dec. 2012, pubmed.ncbi.nlm.nih.gov/23224939/.

Nowakowski, Grzegorz S, et al. “Using Smudge Cells on Routine Blood Smears to Predict Clinical Outcome in Chronic Lymphocytic Leukemia: A Universally Available Prognostic Test.” Mayo Clinic Proceedings, U.S. National Library of Medicine, 8 Apr. 2007, pubmed.ncbi.nlm.nih.gov/17418074/.

Ostojska, Magdalena, et al. “Car-T Cell Therapy in the Treatment of Pediatric Non-Hodgkin Lymphoma.” Journal of Personalized Medicine, U.S. National Library of Medicine, 10 Nov. 2023, pmc.ncbi.nlm.nih.gov/articles/PMC10672004/#B111-jpm-13-01595.

Palma, Marzia, et al. “BTK Inhibitors in Chronic Lymphocytic Leukemia: Biological Activity and Immune Effects.” Frontiers in Immunology, U.S. National Library of Medicine, 1 July 2021, pubmed.ncbi.nlm.nih.gov/34276674/.

Perini, Guilherme Fleury, et al. “Bcl-2 Inhibition as Treatment for Chronic Lymphocytic Leukemia.” Current Treatment Options in Oncology, U.S. National Library of Medicine, 10 June 2021, pubmed.ncbi.nlm.nih.gov/34110507/.

Petros, AM, et al. “Solution Structure of the Antiapoptotic Protein Bcl-2.” Proceedings of the National Academy of Sciences of the United States of America, U.S. National Library of Medicine, 27 Feb. 2001, pubmed.ncbi.nlm.nih.gov/11248023/.

Poeschel, Viola, et al. “Four versus Six Cycles of Chop Chemotherapy in Combination with Six Applications of Rituximab in Patients with Aggressive B-Cell Lymphoma with Favourable Prognosis (Flyer): A Randomised, Phase 3, Non-Inferiority Trial.” Lancet (London, England), U.S. National Library of Medicine, 21 Dec. 2019, pubmed.ncbi.nlm.nih.gov/31868632/.

Prochownik, Edward V., and Huabo Wang. “Lessons in aging from Myc knockout mouse models.” Frontiers in Cell and Developmental Biology, U.S. National Library of Medicine, 9 Aug. 2023, https://pubmed.ncbi.nlm.nih.gov/37621775/.

Rai, K R, et al. “Clinical staging of chronic lymphocytic leukemia.” Blood, U.S. National Library of Medicine, Aug. 1975, https://pubmed.ncbi.nlm.nih.gov/1139039/.

Ramos, Carlos A, et al. “CD19-Car Trials.” Cancer Journal (Sudbury, Mass.), U.S. National Library of Medicine, 2014, pubmed.ncbi.nlm.nih.gov/24667955/.

Riches, John C, and John G Gribben. “Understanding the Immunodeficiency in Chronic Lymphocytic Leukemia: Potential Clinical Implications.” Hematology/Oncology Clinics

of North America, U.S. National Library of Medicine, Jan. 2013, pubmed.ncbi.nlm.nih.gov/23561470/.
Rosales, Tilman, et al. “Partition of Myc into Immobile vs. Mobile Complexes within Nuclei.” Scientific Reports, U.S. National Library of Medicine, 2013, pubmed.ncbi.nlm.nih.gov/23739641/.
Sadelain, Michel. “Car Therapy: The CD19 Paradigm.” The Journal of Clinical Investigation, U.S. National Library of Medicine, 1 Sept. 2015, pubmed.ncbi.nlm.nih.gov/26325036/.
Sangle N, and Nat Pernick. “Anatomy & histology-lymph nodes.” PathologyOutlines.com website, 24 July 2025, https://www.pathologyoutlines.com/topic/lymphnodesanatomy.html.
Schmitt, C A, et al. “INK4A/ARF Mutations Accelerate Lymphomagenesis and Promote Chemoresistance by Disabling P53.” Genes & Development, U.S. National Library of Medicine, 15 Oct. 1999, pubmed.ncbi.nlm.nih.gov/10541553/.
Scott, DW, et al. “Prognostic Significance of Diffuse Large B-Cell Lymphoma Cell of Origin Determined by Digital Gene Expression in Formalin-Fixed Paraffin-Embedded Tissue Biopsies.” Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, U.S. National Library of Medicine, 10 Sept. 2015, pubmed.ncbi.nlm.nih.gov/26240231/.
Sears, R, et al. “Multiple Ras-Dependent Phosphorylation Pathways Regulate MYC Protein Stability.” Genes & Development, U.S. National Library of Medicine, 1 Oct. 2000, pubmed.ncbi.nlm.nih.gov/11018017/.
Sehn, L.H, et al. “Nivolumab for Relapsed/Refractory Classical Hodgkin Lymphoma after Autologous Transplant: Full Results after Extended Follow‐up of the Phase 2 Checkmate 205 Trial - Fanale - 2017 - Hematological Oncology - Wiley Online Library.” Wiley Online Library, 7 June 2017, onlinelibrary.wiley.com/doi/full/10.1002/hon.2437_124.
Singh, Simar Pal, et al. “Role of Bruton’s Tyrosine Kinase in B Cells and Malignancies.” Molecular Cancer, U.S. National Library of Medicine, 19 Feb. 2018, pubmed.ncbi.nlm.nih.gov/29455639/.
SJ, Veis, et al. “BCL-2-Deficient Mice Demonstrate Fulminant Lymphoid Apoptosis, Polycystic Kidneys, and Hypopigmented Hair.” Cell, U.S. National Library of Medicine, 22 Oct. 1993, pubmed.ncbi.nlm.nih.gov/8402909/.
Smith, Victoria M, et al. “Specific Interactions of Bcl-2 Family Proteins Mediate Sensitivity to Bh3-Mimetics in Diffuse Large B-Cell Lymphoma.” Haematologica, U.S. National Library of Medicine, 10 Oct. 2019, pubmed.ncbi.nlm.nih.gov/31601689/.
Tagawa, H, et al. “Genome-Wide Array-Based CGH for Mantle Cell Lymphoma: Identification of Homozygous Deletions of the Proapoptotic Gene Bim.” Oncogene, U.S. National Library of Medicine, 17 Feb. 2005, pubmed.ncbi.nlm.nih.gov/15608680/.
Tari, Kaveh, et al. “Venetoclax Drug Increases the Apoptosis of T and B Acute Lymphoblastic Leukemia Cells by Reducing the Expression of Bcl-2.” International Journal of

Molecular and Cellular Medicine, U.S. National Library of Medicine, 2023, pmc.ncbi.nlm.nih.gov/articles/PMC11092900/.

Thompson, Philip A, et al. “Fludarabine, Cyclophosphamide, and Rituximab Treatment Achieves Long-Term Disease-Free Survival in IGHV-Mutated Chronic Lymphocytic Leukemia.” Blood, U.S. National Library of Medicine, 22 Oct. 2015, pubmed.ncbi.nlm.nih.gov/26492934/.

Vidlarova, Monika, et al. “Recent Advances in Methods for Circulating Tumor Cell Detection.” International Journal of Molecular Sciences, U.S. National Library of Medicine, 15 Feb. 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC9959336/.

Vitolo, Umberto, et al. “Obinutuzumab or Rituximab plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Previously Untreated Diffuse Large B-Cell Lymphoma.” Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, U.S. National Library of Medicine, 10 Aug. 2017, pubmed.ncbi.nlm.nih.gov/28796588/.

Wang, Haolan, et al. “Targeting P53 Pathways: Mechanisms, Structures, and Advances in Therapy.” Signal Transduction and Targeted Therapy, U.S. National Library of Medicine, 1 Mar. 2023, http://pubmed.ncbi.nlm.nih.gov/36859359/

Wang, Zhenguang, et al. “Current Status and Perspectives of Chimeric Antigen Receptor Modified T Cells for Cancer Treatment.” Protein & Cell, U.S. National Library of Medicine, 2 May 2017, pubmed.ncbi.nlm.nih.gov/28466386/.

Welcker, Markus, and Bruce E Clurman. “FBW7 Ubiquitin Ligase: A Tumour Suppressor at the Crossroads of Cell Division, Growth and Differentiation.” Nature Reviews. Cancer, U.S. National Library of Medicine, 8 Feb. 2008, pubmed.ncbi.nlm.nih.gov/18094723/.

Williams, Ashley B., and Bjorn Schumacher. “p53 in the DNA-Damage-Repair Process.” Cold Spring Harbor Perspectives in Medicine, U.S. National Library of Medicine, 2 May 2016, https://pubmed.ncbi.nlm.nih.gov/27048304/.

Xu-Monette, Zijun Y, et al. “Dysfunction of the TP53 Tumor Suppressor Gene in Lymphoid Malignancies.” Blood, U.S. National Library of Medicine, 19 Apr. 2012, pmc.ncbi.nlm.nih.gov/articles/PMC3335376/#sec1.

Xu-Monette, Zijun Y, et al. “Clinical and Biologic Significance of MYC Genetic Mutations in de Novo Diffuse Large B-Cell Lymphoma.” Clinical Cancer Research : An Official Journal of the American Association for Cancer Research, U.S. National Library of Medicine, 15 July 2016, pmc.ncbi.nlm.nih.gov/articles/PMC4947447/#SD1.

Yoo, Kwai Han. “Staging and response assessment of lymphoma: a brief review of the Lugano classification and the role of FDG-PET/CT.” Blood Research, U.S. National Library of Medicine, 30 Apr. 2022, https://pmc.ncbi.nlm.nih.gov/articles/PMC9057662/.

.

The Effect of New Metaphysical Knowledge By Leonardo Wang

Abstract

This essay explores how metaphysics can influence a person's values and moral beliefs. By analyzing the three branches of metaphysics: ontology, theology, and temporality, I show how breaking down the foundations of reality can make someone think a lot differently about their previously held morals. I draw from historical and contemporary philosophers, such as Aristotle, Immanuel Kant, as well as modern theorists like Martin Heidegger and Helge Høibraaten. Their work argues that new metaphysical knowledge allows people to re-evaluate what they previously thought about the world's moral systems. Analyzing each metaphysical perspective allows us to gain a much deeper understanding of morals and ethics.

Introduction

People tend to believe that their values and beliefs are absolute and unchangeable. However, metaphysics completely skews this idea. The basic structure of reality is so complex that humans have to break it down in order to understand it. Whether it is viewing time's linearity or taking a humanistic approach on life, one will inevitably question one's moral value when looking at metaphysics. There are three branches of metaphysics once it is broken down: ontology (the study of being), theology (the study of God), and temporality (the study of time). Ontology has two perspectives that can shape one's values and beliefs, humanism and post-humanism. They are total opposites of each other and challenge the way people think about human importance. The second branch of metaphysics studies the relationship of God and religion to metaphysics. The two perspectives of secularism and religion have provided a way to new morals. The third branch of metaphysics is temporality. The nature of time is very subjective, and can be viewed in a multitude of ways, whether it is through linearity or timelessness. As a result of these three arguments, new metaphysical knowledge can have a great impact on one's values and beliefs.

Humanism vs. Post-Humanism

To understand how metaphysics can change established values and beliefs, it is important to start with ontology, in which lies the two perspectives of humanism and post-humanism. The perspective of humanism is quite controversial, as it puts humans as the most important beings in existence. This perspective is most commonly found in older philosophers, who do not recognize that non-human life is quite important in human life. As German philosopher Immanuel Kant once said, "Act so that you treat humanity, whether in your own person or in that of another, always as an end and never merely as a means." Kant's quote here implies that each human has equal worth, so humans should not be treated as tools. Kant places humans at the center of ethical concern and fails to mention how non-human life also has worth. On the other side of the spectrum is post-humanism, which challenges that non-human life has the same worth as human life. This perspective is more commonly found in modern philosophers, who first-handedly know

that non-human life plays a big part in human life, whether through artificial intelligence or animal life. Among them is German philosopher Martin Heidegger. He famously writes in his *Letter on Humanism*, "The human being is not the lord of beings. The human being is the shepherd of being." Heidegger criticized humanism by saying that humans should not be in control but rather take care of being. A lot of modern day philosophers share Heidegger's beliefs that humans are not rulers but caretakers for life. Ultimately, these two ontological perspectives are polar opposites of each other. Whether humans truly are superior to non-human life or not, one must note that these two perspectives can completely change one's values and beliefs.

Secularism vs. Religion

As post-humanism challenges the idea that humans are the superior beings in existence, the question of the existence of God or a higher power arises. At the center of this question is theology. Again, two perspectives emerge. Secularism, which relies on reason and evidence to establish values, separates itself from religion. One philosophical view related to secularism is agnosticism, the belief that the existence or non-existence of God is unknowable. Norwegian philosopher Helge Høibraaten aligns himself with the agnostic view. He writes, "In a postsecular society, the agnostic philosopher can no longer simply ignore the fact that religious traditions continue to articulate moral intuitions that remain inaccessible to secular language alone." Høibraaten argued that secularism is the translation of religion, not elimination. However, in order for secularism and religion to coexist, agnostic philosophers must realize that some religious traditions cannot be translated by secular language. As Høibraaten believed, secularism must coexist with religion in order to have a fair society. This leads into the other perspective of theology: religion. Religion differs from secularism in that it relies on God and divinity to establish metaphysical truths. There are many different religions, each containing different practices. However, one commonality remains: the belief in a higher power or God. Immanuel Kant wrote in *Critique of Practical Reason*, "Two things fill the mind with ever new and increasing admiration and awe... the starry heavens above me and the moral law within me." Kant famously believed that there are two metaphysical realities: natural law and moral law. He points to the moral law as being so powerful that it required the existence of God and the immortality of the soul. To conclude, secularism and religion both provide insight towards the existence of God and how it affects beliefs, and need to coexist in order for an equal society.

Linearity vs. Non-Linearity

While theology inspects how religion and belief in God shape moral beliefs, another overlooked topic in metaphysics is temporality, the study of the nature of time. Time is a subjective and complex concept when viewed through the lenses of metaphysics. As with ontology and theology, two perspectives appear in temporality. The first one says that time is linear. Linear time is how people usually perceive time. It involves time moving in a straight path consisting of the past, the present, and the future. Authors Taylor Cyr and Neal Tognazzini argue,

“Our practices of blaming and praising people for what they’ve done presuppose a temporally ordered world in which actions precede outcomes.” Cyr and Tognazzini assume that if time is linear, people’s moral practices like punishing or complimenting move in a forward direction. Someone does an action, that action leads to consequences, and people judge them based on the consequences. Linear time is built into how humans understand right or wrong. However, questions also emerge from linear time. Suppose time were non-linear. Can one be blamed for the actions that have “yet” to happen? The non-linear perspective of time suggests that the past, present, and future could be repeated, loop or even coexist. An example of a non-linear concept is time travel. Cyr and Tognazzini first assumed that traditional moral systems are based on a linear timeline, but they further expand on that argument by introducing causal loops, a theory of time travel. If a person travels back in time and influences their younger self to do something, were they really in control? The theory of time travel brings a whole new component into how humans view morals in time. Truly, the perception of the linearity of time gives an unique idea about how one’s morals are shaped.

Conclusion

In the end, metaphysics gives humans new ways to think about already established values and beliefs. Many might think that their beliefs can never change from metaphysical knowledge. However, ontology, theology, and temporality provide very different perspectives on how different people’s lives are shaped. The ontological perspective of post-humanism forces us to consider that non-human life is now a big factor in human life, so humans should aim to take care and protect life. The theological perspectives of secularism and religion should coexist for society to be equal, rather than conflict with each other. It makes humans completely reconsider their beliefs when brought up. And finally, temporality brings in the factor of linearity of time, which makes us think about morals and how they might play a part in one’s life through time. Together, these three branches of metaphysics can completely establish our values and beliefs, letting us understand our existence through multiple perspectives. Even if our values and beliefs do not change, learning metaphysical knowledge through each perspective helps us understand the real nature of reality.

Works Cited

Askheim, C., Engebretsen, E., & Haldar, M. (2024). Living happily alone in Plato's Cave? on loneliness, technology and the metaphysics of presence. *Medical Humanities*, *50*(3), 561–569. https://doi.org/10.1136/medhum-2024-012965

Cyr, T., & Tognazzini, N. (2024). What time travel teaches us about moral responsibility. *Journal of Ethics and Social Philosophy*, *26*(3). https://doi.org/10.26556/jesp.v26i3.1770

Høibraaten, H. (2017). 40. religion, metaphysics, freedom: "Faith and knowledge" (2001). *The Habermas Handbook*, 444–460. https://doi.org/10.7312/brun16642-042

Jiang, J. (2024). Secular procreation: Metaphysics of birth in the 21st Century. *Anthropology Today*, *40*(6), 3–6. https://doi.org/10.1111/1467-8322.12924

Macdonald, C. (2005). *Varieties of things: Foundations of contemporary metaphysics*. Blackwell Pub.

Rae, G. (2013). Heidegger's influence on posthumanism. *History of the Human Sciences*, *27*(1), 51–69. https://doi.org/10.1177/0952695113500973

Commuting Time Matters: Evaluating the Effect of Distance and Airfare in Metropolitan Areas By Aarav Bolar

Abstract

This paper investigates the relationship between distance from city center (DFCC) and airfare. Airfare is extremely volatile and dependent on many factors. Although some of the factors have been greatly researched, there are hundreds of factors that still haven't been researched. Fortunately, most researchers have researched extensively on the effect of airfare through many different factors, such as time of booking, day of departure, competition, and many more factors. Nevertheless, the effect of many other factors such as DFCC hasn't been greatly researched.

My research extends on previous research that has been conducted by narrowing parameters, differing from other researchers, to focus on metropolitan areas and the United States. Additionally, I also conducted a methodological gap by focusing on a correlational methodology between the two variables. My methodology involved analyzing over thirty routes per airport, repeated thirty times across three different groups. For each airport, the airfare data from these thirty routes was averaged, and the resulting average fare was used in the scatterplot to compare with the DFCC.

The relationship between DFCC and airfare can aid consumers find cheap airfare without spending hours to do so. Moreover, airlines would be able to identify target markets in order to market cheaper airfare to attract consumer interest. Finally, other researchers would be able to focus on other factors as a focus of research in order to discover what airfare is really dependent on.

After analysis on scatterplots from three different groups: short haul routes, mid haul routes, and long haul routes, the long haul routes had the greatest correlation between DFCC and airfare. DFCC and airfare also had a positive correlation, so as the DFCC increased for the airport, the cost of the airfare resulted in an increase as well.

Introduction

When booking air travel within the United States, passengers focus on a variety of different factors before confirming a ticket, such as the price of the ticket, flight duration, airline reputation, and many other factors. Many routes across the United States have multiple airlines that serve the route, so passengers have to take this into consideration as well. However, at times passengers, especially those who live in large metropolitan areas, also have multiple airports to decide from. With many metropolitan areas served by multiple airports, travellers have a choice between flying through a more convenient, closer airport, or flying from a further airport that may be cheaper or possibly more expensive.

The economic and logistical dynamics behind airfare pricing are complex, shaped by many factors such as airport demand, operating costs, competition among airlines, and passenger convenience. Airports closer to the downtown areas may charge higher landing fees due to the

higher infrastructure costs and the limited space, as the city is more compact. However, many budget airlines choose airports that are farther from urban centers, which could potentially result in cheaper fares. Ultimately, according to John Bitzan, PhD in Economics and professor at North Dakota State University, smaller regional airports don't have nearly as much competition as larger airports, causing airfare in smaller airports to be greater than larger airports (Bitzan et al.). Consequently, smaller airports may have cheaper airfare than larger airports, but sometimes this may not hold true in all cases.

A study conducted by Woohyun Cho, PhD in Department of Management in the University of New Orleans, titled "Route competition in multi-airport cities: an analysis of US air fares" examines the passenger demand and the airfare in multi-airport cities (Cho, et al.). In addition to this study, another study organized by T. Godwin, titled "An Empirical Analysis of Delhi - Mumbai Sector Flight Fares" analyzes the main factors that affect airfare between two of the most popular metropolitan areas in India (Godwin). These are key papers that shaped the direction of the paper and what served as a foundation of my paper.

Literature Review

Airfare is the price of an airline ticket at booking, whether that is through an airline's website or a third party source such as Kayak or Expedia. Airfare is very mercurial depending on the time of travel, distance of travel, and the time of booking. According to expert Greg Iacurci, reporter of personal finance at CNBC, airfare to countries such as Tokyo, Canada, the Middle East, and South America have experienced jumps in prices, but cities in Europe are experiencing a decrease in prices (Iacurci). Another source by Andy Smith, certified financial programmer for 35+ years, adds on to this thought but also informs consumers that the ideal time to book flight tickets is approximately 21 to 74 days before departure. They also advise their consumers to book holiday travel more in advance as compared to other times of travel (Smith et al.). Airfare fluctuations and cheaper prices were all caused by the deregulation movement.

According to Steven Morrison, professor of economics at Northeastern University, the deregulation event was a time in 1978 when federal control was removed from various areas of the airline industry, such as fares, routes, and market entry for new airlines, allowing the market to open up to airlines to set their own fares and routes without strict government interference. Since the deregulation event of the late 20th century, some researchers claim that the prices of airfare have dropped considerably (Morrison, et al.). Morrison also claims that airfare has had a massive drop of 33% between 1976 and 1993, indicating a significant drop in the airfare prices worldwide, and the effect that deregulation provided airlines to have better service onboard their aircraft and increase frequency between two cities. The more frequency a route has, the more convenience it has for airlines, and the more demand that exists between these two cities (Morrison, et al.). To add on to this thought, PhD candidate Qi Zhang from the School of Geography and Planning in Sun Yat-sen University, Guangzhou discusses how deregulation has led to a growth in the market between China and Oceania. Deregulation played a role in the increase of route frequency between China and Oceania, bringing increased competition between

the two regions and decreased airfare (Zhang, et al.). Although this work differs from the previous work from Morrison by examining a different region, they both back up the same argument by analyzing different regions. In contrast, Jutta Joesch, assistant professor in the University of Utah over Consumer Studies, claims that deregulation led to unintended consequences, such as a later increase in airfares again, reduced competition, and a decline in service quality, baggage handling, and the on-time performance of airlines (Joesch, et al.). Deregulation has benefited the economy and the consumer in terms of airfare, but also has its own consequences on airfare and other unintended consequences.

Moreover, there are many major factors that affect airfare. A key factor is competition, also known by the competitive prices that airlines put on the consumers so that consumers choose their airline. These are mainly influenced by airline alliances, when airlines join together to allow for more connectivity between two countries. In the SkyTeam alliance, one of the three major airline alliances in the world, Delta and Air France have created a tight alliance in order to boost connectivity between European or American destinations that one doesn't fly to. According to Xiang Wan, studied business at Robert H. Smith School of Business in the University of Maryland, airlines in an alliance perform a codeshare agreement, where they can book flights on a different airline with the other airline, hence decreasing prices on airfare. However, as alliances increase, the price competition also increases (Wan). It's not just about other airlines, though; other methods of transportation can also have an affect on airfare prices. For example, Peiwen Zhang, professor at the School of Economics and Management in Civil Aviation Flight in University of China, Guangzhou found that in China, the high speed railroad can be quicker for passengers than air travel, depending on the passenger's distance from the city center, which would variate the success of a specific airline route (Zhang, et al.). Although airlines are huge competitors of each other, other developing forms of transportation such as the high speed rail can also affect airfare.

Most research up to this point has been composed of airfare and competition of airfare in general. A few researchers have discussed specific countries or areas, but most consider a different region or methodology or approach the relationship using different parameters. My research builds upon existing research by filling in a methodological gap, examining the relationship between DFCC and airfare, with a primary focus on large metropolitan areas in the United States, while also using a different methodology. In addition, my study includes nearby airports that, although located outside the metropolitan area, are frequently used by passengers traveling to major cities.

By understanding the relationship between these two cities, other giant and growing metropolises can apply this knowledge into their city, designing the city to be more convenient for passengers too far from the other airport, and to have more competitive fares. Airlines can also use this information while planning more routes in between two major metropolises, in order to determine the most advantageous way in which they could reduce prices to effectively attract consumers in the market. Consumers seeking for lower airfare can also use information to their advantage in finding cheaper flights for their next business or leisure trip in order to avoid

paying ludicrously high prices. This leads to the question, “To what extent does the distance between an airport and the city center impact airfare pricing in the United States?”

Methodology

In order to investigate the correlation between the distance between an airport and the city center (DFCC) and the airfare, I conducted a correlational methodology. A correlational study is essentially using two quantitative variables to determine whether there is a correlation, without implying causation. By employing this methodology, it would aid me in the assistance of comparing two different quantitative variables. The primary goal of my research is to determine whether there is any effect on airfare from the DFCC and airfare; hence, a correlational methodology would logically lead me to approach at a conclusion that would be most appropriate in answering the research question. Additionally, implementing a correlational methodology would be the best fit for this topic, because a correlational methodology grants me to use quantitative variables without the necessity to manipulate variables.

Other methodologies could have been suitable for my research, for instance an experimental design. Nevertheless, this design wouldn’t have worked the best, because although the DFCC changes throughout different metropolitan areas, the DFCC can’t be manually altered, which deviates from the aim of an experimental procedure. Using a correlational study allows me to determine a correlation using the DFCC and the airfare that has already been presented, without the need to modify any pre-existing data.

My methodology is outlined in a few basic steps. Before understanding the methodology, there were three different subgroups within the main methodology: Group A, Group B, and Group C. I decided to use three different groups, because three trials are best practices in order to check for consistency among data. Group A focuses on short-haul flights between 200-600 miles of distance between each other, Group B focuses on mid-haul flights between 900-1400 miles, and Group C measures long-haul flights between 1600-2500 miles.

The first step of my methodology was to research all the possible routes that I could investigate. I wanted to use an array of 5-7 different routes per airport, however that isn’t possible for some of the smaller airports that I am investigating. Therefore, it proved sometimes difficult to adhere to my guideline at all times, so at times I investigated only 3 unique routes. This step of my methodology proved beneficial, because this set as a foundation towards finding the overall data. I found 30 routes per airport, some reused and some unique, because having a sample trial of 30 is ideal for statistical measurement, in order to avoid the interference of outliers.

After this step was complete, the second step of my methodology was to assign random dates for all the routes I found per airport. I randomly generated numbers between 1 and 120, because there are 120 days between February 1st and June 1st; each number represents how many total days after February 1st. I decided to look between the dates of February 1, 2025 and June 1, 2025, because if I looked for flights too far or close in advance, the airfare may greatly increase due to the proximity of the date. Additionally, according to Terry Lane, prices during the

summer tend to be higher than normal (Lane). This would potentially cause outliers or inconsistent data, having the end of the date range to be around June 1st would be ideal to avoid possible outliers.

Next, I researched any major events that may increase airfare, because any major event would result in an anomaly in the surrounding day's airfare. Hence, I randomly assigned another date to this route and any routes with a major event interfering with the route I am researching, in order to prevent outliers.

After that, I was ready to start researching the actual airfare. Using the dates that I randomly assigned and the routes previously researched, I used a major travel network site such as Kayak or Expedia in order to find the lowest airfare on any airline on that specific date because these travel websites are helpful in showing the cheapest airfare. I also made sure to think in the eyes of the consumer, and only commit to routes that made sense. Specifically, I made sure to ignore all routes with a 5 hour or longer layover or routes with a stopover hundreds of miles out of the way in either of the airports, because the average consumer looks for cheap flights that are also relatively convenient. I researched all 30 airfares, and calculated the averages of each of these airfares, which will come in handy when I conduct the correlational test. I repeated this step for all 30 airports in each of the three groups until I had sufficient data that I could use to analyze the data.

In order to analyze data, it was first crucial to determine the DFCC, as I had not previously researched until I finished calculating the averages. In order to calculate the DFCC, I used the ground distance between the airport and the town hall. I decided to use the town hall, because according to Matthew J. Holian, economics professor at San Jose State University, the city center can be measured from the city hall (Holian). Using these ground distances were most efficient, because most people travel through vehicle transportation and not air distance to the airport. Using the DFCC on the x-axis, and the average airfare on the y-axis, I created a scatterplot for each of the three groups. From there, I drew a trendline to determine what happens to the airfare as the DFCC increases, which wasn't necessary to reach my conclusion.

In order to prove an actual correlation, more statistical values such as the r-value needed to be calculated for each of the scatterplots. I decided to use the r-value to determine a correlation, because using this value would best prove and measure the correlation between the two variables. I calculated the r-value for each of the three scatterplots, and analyzed my data based on this value.

Although using my methodology would have the most benefits in achieving results, there are also some limitations with this methodology overall. Specifically, there are only a limited number of airports that would fall into a multi-airport metropolitan area, proving difficult to attain a sample size that would be large enough to fully eliminate margin of error. In addition, there are a lot of confounding variables. Airfare is very volatile, dependent on many variables that may affect pricing on one day over another. Having a lot of confounding variables, such as a lack of sample size and unpredictability of airfare, would make me need to observe and eliminate the confounding variables that may interrupt the conclusions of my study. While remaining

vigilant of these limitations, my conducted study would be an ideal way to determine a correlation between DFCC and airfare.

Results

Based on the methodology conducted, in order to arrive at a conclusion whether the distance from city center to airport (DFCC) affects the airfare, I collected these results from this methodology. The results represent two different versions of the data collected: firstly, a table version to provide visual context of the specific data collected, and secondly, the scatterplot with a trendline to determine vital statistical values. I collected the results in three different groups: Group A (defined as short-haul routes), Group B (defined as mid-haul routes), and Group C (defined as long-haul routes). By doing so, each of these groups provided a different experimental trial while being able to analyze different data sets, in order to ensure consistency between groups and data.

Figure 1 provides the average data values that I calculated (Expedia). In the table below, Figure 1 provides the airport IATA code (a three-letter code that is commonly used to refer to airports), the greater metropolitan area associated with this airport, the DFCC, and the average airfare.

The data in the IATA code, metropolitan area, and the DFCC columns will remain constant throughout Figure 1, Figure 3, and Figure 5, because I didn't change the airports I was looking at throughout the study; I only changed the routes I was looking at. The DFCC represents the distance through road systems from the city's town hall and airport on the fastest route, and the average airfare is the average airfare calculated from a random sample of 30 different routes (reference sample tables in Appendix). The DFCC is rounded to one decimal place, and the average airfare is rounded to two decimal places. The average airfare was collected from Expedia flight data, throughout Figures 1 through 6. Before viewing the data, it is important to note that each of the data tables are not in order of DFCC, and that they are in the order that I collected data.

IATA Code	Metropolitan Area	DFCC	Average Airfare
JFK	New York	17.6 mi	$108.23
LGA	New York	13.5 mi	$126.90
EWR	New York	12.4 mi	$105.23
IAD	Washington	23.6 mi	$163.50
DCA	Washington	4.1 mi	$135.87
BWI	Washington	30.5 mi	$136.83
LAX	Los Angeles	18.3 mi	$59.77

ONT	Los Angeles	38.9 mi	$109.40
SNA	Los Angeles	38.2 mi	$143.50
BUR	Los Angeles	13.4 mi	$113.07
LGB	Los Angeles	25.1 mi	$129.27
SFO	Bay Area	35.2 mi	$63.20
SJC	Bay Area	3.2 mi	$100.97
OAK	Bay Area	34.5 mi	$98.93
MRY	Bay Area	71.8 mi	$190.27
STS	Bay Area	106.0 mi	$135.57
ISP	New York	54.0 mi	$212.87
MIA	Miami	7.4 mi	$171.37
FLL	Miami	26.8 mi	$80.60
PBI	Miami	69.4 mi	$199.60
HPN	New York	36.4 mi	$300.27
BOS	Boston	3.2 mi	$92.30
MHT	Boston	53.0 mi	$302.00
ORH	Boston	52.9 mi	$240.73
ORD	Chicago	16.0 mi	$111.73
MDW	Chicago	10.5 mi	$99.67
DFW	Dallas	16.5 mi	$149.87
DAL	Dallas	5.6 mi	$153.23
IAH	Houston	15.8 mi	$146.03
HOU	Houston	9.2 mi	$156.53

Fig 1: Table of airports and average short-haul route airfare

This table provides the data points that will be reflected on the scatterplot, as shown in

Figure 2 below. Figure 2 provides the points in a scatter plot context, where the DFCC is plotted on the x-axis, and the average airfare is plotted on the y-axis. A trendline is also provided on the graph, in order to provide a direction of correlation.

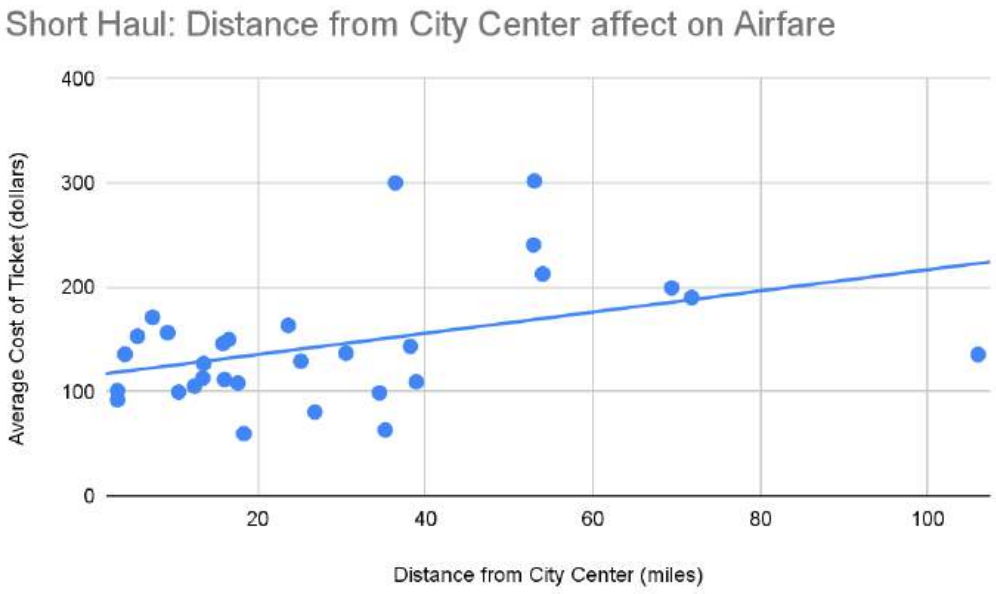


Fig 2: Scatterplot of airports and average short-haul route airfare

From the scatterplot, I calculated statistical values such as slope (m) and r-value (r). Other values, such as the y-intercept and r^2-value could've been easily calculated from the scatterplot, however these values proved unnecessary in the broader context of my study. From Figure 2, the slope is approximately 1.015, and the r-value, from Figure 2 is calculated to be approximately 0.409. Although this provides some indication of correlation, an r-value of 0.409 provides evidence of a relatively weak to medium linear correlation.

Figure 3 represents the average data values that were calculated, however for medium haul routes this time (Expedia). Medium haul routes were defined as routes that were between 900-1400 miles away from the airport of origin.

IATA Code	Metropolitan Area	DFCC	Average Airfare
JFK	New York	17.6 mi	$92.23
LGA	New York	13.5 mi	$101.77
EWR	New York	12.4 mi	$105.90
IAD	Washington	23.6 mi	$156.07
DCA	Washington	4.1 mi	$161.53

Fig 3: Table of airports and average mid-haul route airfare

The table provides the data points that will be shown on a scatter plot in Figure 4. Figure

4 provides the data presented in Figure 3 on a scatter plot, similar to Figure 2 however the data are determined through mid-haul routes.

Figure 4

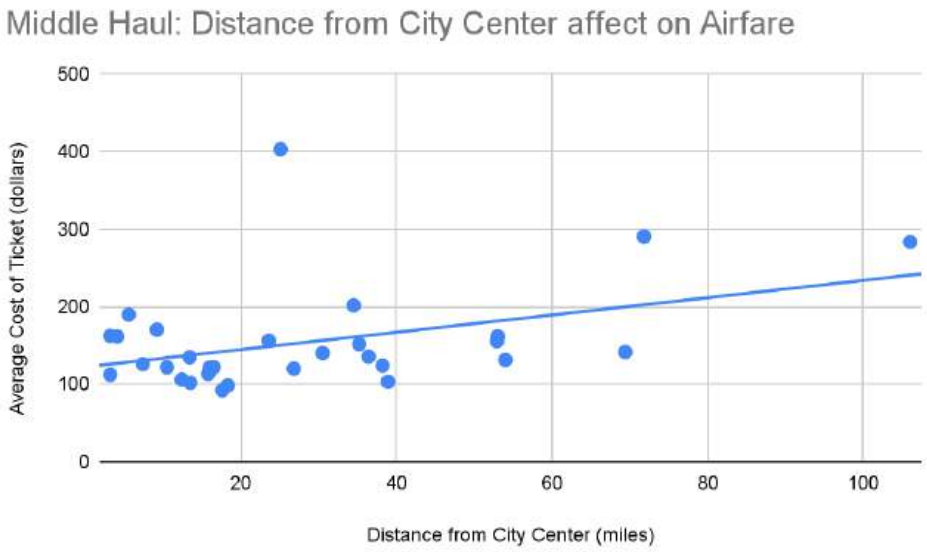


Fig 4: Scatterplot of airports and average mid-haul route airfare

From the scatterplot above, pivotal statistical values were calculated, such as the slope, y-intercept, r-value, and r^2-value. The slope value from Figure 4 was derived to be 1.115 dollars per mile. However, the r-value in Figure 4 is 0.405, computed to be approximately the same as Figure 2. In other words, the correlation between DFCC and airfare for mid-haul routes is relatively weak to moderate. Although the trendline looks more organized, the statistical values still haven't changed, implying that there is no real difference in the relationship.

Figure 5 represents Group C of the data collected, primarily the long-haul routes of the study (Expedia). The long-haul routes are defined as routes between 1600-2500 miles from the origin airport.

IATA Code	Metropolitan Area	DFCC	Average Airfare
JFK	New York	17.6 mi	$163.03
LGA	New York	13.5 mi	$147.53
EWR	New York	12.4 mi	$165.27
IAD	Washington	23.6 mi	$194.00
DCA	Washington	4.1 mi	$172.07

BWI	Washington	30.5 mi	$133.57
LAX	Los Angeles	18.3 mi	$112.50
ONT	Los Angeles	38.9 mi	$150.70
SNA	Los Angeles	38.2 mi	$230.47
BUR	Los Angeles	13.4 mi	$249.23
LGB	Los Angeles	25.1 mi	$288.27
SFO	Bay Area	35.2 mi	$155.97
SJC	Bay Area	3.2 mi	$159.10
OAK	Bay Area	34.5 mi	$226.57
MRY	Bay Area	71.8 mi	$222.00
STS	Bay Area	106.0 mi	$271.77
ISP	New York	54.0 mi	$277.83
MIA	Miami	7.4 mi	$169.23
FLL	Miami	26.8 mi	$185.03
PBI	Miami	69.4 mi	$191.90
HPN	New York	36.4 mi	$218.03
BOS	Boston	3.2 mi	$152.40
MHT	Boston	53.0 mi	$224.33
ORH	Boston	52.9 mi	$199.73
ORD	Chicago	16.0 mi	$132.60
MDW	Chicago	10.5 mi	$177.53
DFW	Dallas	16.5 mi	$119.13
DAL	Dallas	5.6 mi	$266.73
IAH	Houston	15.8 mi	$113.37
HOU	Houston	9.2 mi	$109.00

Fig 5: Table of airports and average long-haul route airfare

This table contains the statistical values that are going to be represented on the scatterplot in Figure 6. The scatter plot in Figure 6 represents the relationship between DFCC and airfare in the context of long-haul routes, similar to Figures 2 and 4.

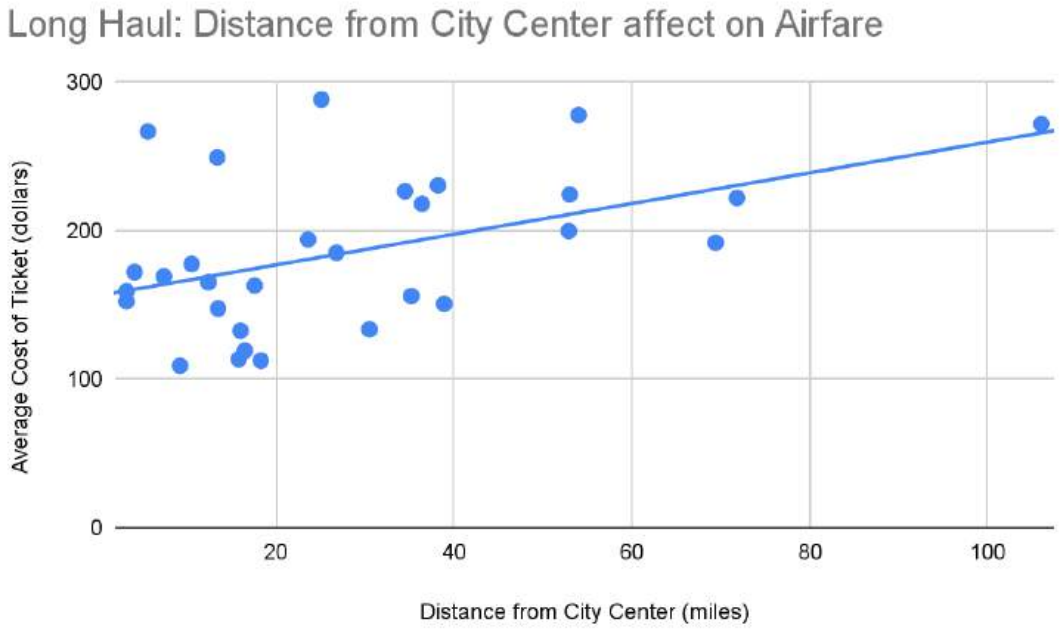


Fig 6: Scatterplot of airports and average long-haul route airfare

From the scatter plot in Figure 6, the slope of the trendline in Figure 6 is found to be 1.034 dollars per mile, which is a slight decrease from the slopes found in Figures 2 and 4. The r-value of Figure 6 is 0.478, showing a moderate correlation between DFCC and airfare for long-haul routes. Thus, the r^2-value in Figure 6 increased considerably to 0.229, although not showing a strong correlation, showing that these values have slightly increased as compared to short-haul and mid-haul routes.

Discussions/Analysis

Based on the research question "To what extent does the distance between an airport and the city center impact airfare pricing in the United States," and the results discussed previously, I came to the following conclusions. The most significant statistical value that I determined from my scatterplot was the r-value, which proved extremely crucial in the overall analysis of my study. The r-value was used to determine the strength of the correlation between distance from city center to the airport (DFCC) and airfare, which precisely answers the question at hand.

According to Figure 2 (Group A), the r-value was determined to be 0.409, which can indicate a relatively weak correlation. However, we shouldn't conclude that there is no correlation due to a weak correlation; the correlation still means something since airfare can be affected through a variety of factors. Hence, the r-value means in my study that a greater distance

can modify the airfare, but there are a lot of other factors that can result in lower airfares or higher airfares. Additionally, the slope of Figure 2 is determined to be 1.015, indicating as the DFCC goes up, airfare has a positive correlation and will increase as well.

Similarly, according to Figure 4 (Group B), the r-value was slightly lower, determined to be 0.405. The correlation determines a similar trend to Group A: there is a little correlation that can affect airfare but not exactly a definitive relationship. The slope of Figure 4 is approximately 1.115, which indicates a positive correlation between DFCC and airfare. Moreover, there is a slightly steeper slope than in Group A, indicating a slightly greater difference between DFCC and airfare.

Finally, according to Figure 6 (long-haul routes), the r-value is higher, computed at approximately 0.478, also indicating a low correlation between DFCC and airfare. Compared to Group A and B, Group C has a slightly higher correlation, meaning DFCC impacted airfare a slightly higher amount compared to Groups A and B. Additionally, the slope of Figure 6 is 1.034, suggesting that a higher DFCC would lead to a greater increase in airfare compared to Group A, but a smaller increase than in Group B. The slope also indicates a positive correlation between DFCC and airfare, consistent with Groups A and B.

Overall, given the results from Group A, B, and C, a general conclusion between DFCC and airfare from my study can be summarized that there is a relatively weak correlation between DFCC and airfare. However, the relationship can still be considered as a possible factor of airfare due to its linear relationship, and the status of airfare being dependent on multiple factors. DFCC can be considered a less significant factor, while capable of influencing airfare, wouldn't lead to any substantial changes. Additionally, as DFCC increases from the city center, the average airfare in the airport would tend to increase.

In addition, another study conducted by Woohyun Cho concluded that the DFCC does have a correlation with airfare (Cho, et al.). Cho's study aligns with mine more closely, as they used equations and quantitative measures, similar to my study. My study is consistent with the idea that there is some relationship between DFCC and airfare; both of our studies indicate that DFCC can't be ignored as a factor of airfare. However, my presented conclusion isn't persistent with the idea that there is a great influence between DFCC and airfare, as my research determined a weaker correlation than Cho's, revealing that the correlation between DFCC and airfare may not be as significant as previously thought.

My research can aid a lot of consumers and airlines. My findings can aid customers with the relationship between DFCC and airfare and how much DFCC affects airfare. Since consumers strive to find cheaper airfare through a variety of methods, this relationship can better aid consumers have an understanding of where to look for cheaper flights. For example, using the results from my research, consumers would tend to search more closely for flights in Washington Reagan Airport (DCA) than Washington Dulles Airport (IAD) because according to my research, an airport with a closer DFCC would tend to have cheaper airfare. Although the correlation isn't strong, this relationship could still help consumers identify cheaper airfare.

However, my research can only be limited to the United States, because I only analyzed major airports in the United States. Therefore, my study won't apply to other nations, such as Mexico, Brazil, or Japan. This is a further direction that other researchers could take; as the correlation between DFCC and airfare is more clear in the United States, other researchers could do a similar study in other countries, or even take a look in the global perspective.

Next, my research can also benefit airlines by helping them discount flights towards their advantage. Since a higher DFCC distance generally results in a higher airfare, airlines could benefit from this relationship by discounting airfare that is further from the DFCC. Using this strategy, consumers who regularly travel from the further airport would find this price, thus resulting in immediate attention for consumers. This would result in a greater profit for airlines overall.

However, my study only applies to a metropolitan area, and not the entire state as a whole. Airports that are used by the metropolitan area that is far away from the metropolitan area could qualify as a plausible alternative, but my study doesn't apply to the entire state, therefore this makes the risk of airlines misinterpreting an airport to be part of a metropolitan area when it isn't. Other researchers can take my research even further by analyzing how the relationship applies to the entire state, rather than focusing on just specific to the metropolitan area. This would give consumers a general direction of how this relationship can apply to not just the metropolitan area airports, but the entire state as a whole.

Conclusion

The relationship between DFCC and airfare are known to be directly related, however airfare isn't exactly a great factor of DFCC. Although DFCC can be considered a factor of airfare, there isn't necessarily a strong relationship between the two variables. Other factors, such as time of booking and competition with other airlines show more correlation between each other, and effectively demonstrate to consumers what to look for in order to find cheap airfare. By determining the factors that make up and fluctuate airfare, consumers would be able to effortlessly find cheap airfare without wasting hours trying to do so.

In addition, I determined that there is the strongest relationship between DFCC and airfare in long-haul flights, and the greatest effect of DFCC on mid-haul flights. People flying approximately 900-1400 miles should consider looking for airfare closer to the city center in order to find the cheapest deal. Subsequently, passengers travelling between 1600-2500 miles away should factor in DFCC before booking, because long-haul flights have the most correlation between DFCC and airfare.

Works Cited

Bitzan, John, and Junwook Chi. “Higher Airfares to Small and Medium Sized Communities - Cost.” Journal of Transport Economics and Policy, University of Bath, 1 Jan. 1970, ideas.repec.org/a/tpe/jtecpo/v40y2006i3p473-501.html. Accessed 02 Aug. 2025

Cho, Woohyun, et al. “Route Competition in Multi-Airport Cities: An Analysis of US Air Fares.” Route Competition in Multi-Airport Cities: An Analysis of US Air Fares, July 2012, www.researchgate.net/publication/236772541_Route_Competition_in_Multi-Airport_Cities_An_Analysis_of_US_Air_Fares. Accessed 13 Mar. 2025

dell’Olio, Luigi, et al. “Modeling Airport Choice Using a Latent Class Logit Model.” MDPI, Multidisciplinary Digital Publishing Institute, 10 Aug. 2023, www.mdpi.com/2226-4310/10/8/703. Accessed 13 Mar. 2025

Dhir, Rajeev, et al. “The Best Days to Book the Cheapest Flights.” Investopedia, Investopedia, 7 Nov. 2024, www.investopedia.com/ask/answers/113015/do-plane-tickets-get-cheaper-closer-date-departure.asp. Accessed 13 Mar. 2025

“Expedia Travel: Vacation Homes, Hotels, Car Rentals, Flights & More.” Expedia, 1996, www.expedia.com/. Accessed 13 Mar. 2025

Godwin, T. “An Empirical Analysis of Delhi - Mumbai Sector Flight Fares.” International Journal of Business Analytics (IJBAN), IGI Global Scientific Publishing, 1 Jan. 2017, www.igi-global.com/gateway/article/187209. Accessed 13 Mar. 2025

Holian, Matthew J. “Where Is the City’s Center? Five Measures of Central ...” Where Is the City’s Center? Five Measures of Central Location, Jan. 2019, www.huduser.gov/portal/periodicals/cityscpe/vol21num2/ch12.pdf. Accessed 13 Mar. 2025

Iacurci, Greg. “Flying Is Cheaper in 2024. But Not for Some Destinations.” CNBC, CNBC, 20 May 2024, www.cnbc.com/2024/05/20/flying-is-cheaper-in-2024-but-not-for-some-destinations.html. Accessed 13 Mar. 2025.

Joesch, Jutta M, and Cathleen D. Zick. “Evidence of Changing Contestability in Commercial Airline Markets during the 1980s - Joesch - 1994 - Journal of Consumer Affairs - Wiley Online Library.” Evidence of Changing Contestability in Commercial Airline Markets During the 1980s, Summer 1994, onlinelibrary.wiley.com/doi/pdf/10.1111/j.1745-6606.1994.tb00812.x. Accessed 13 Mar. 2025

Lane, Terry. “Summer Flyers Face Higher Airfares, Crowded Planes.” Investopedia, Investopedia, 5 Apr. 2024, www.investopedia.com/summer-flyers-face-higher-airfares-crowded-planes-8622448?utm_source=chatgpt.com. Accessed 02 Aug. 2025

Winston, Clifford, and Steven A. Morrison. “The Fare Skies: Air Transportation and Middle America.” Brookings, 1 Sept. 1997,

www.brookings.edu/articles/the-fare-skies-air-transportation-and-middle-america/. Accessed 13 Mar. 2025.

Wan, Xiang, et al. “Assessing the Price Effects of Airline Alliances on Parallel Routes.” Transportation Research Part E: Logistics and Transportation Review, Pergamon, 5 Mar. 2009, www.sciencedirect.com/science/article/abs/pii/S1366554509000064?via%3Dihub. Accessed 13 Mar. 2025.

Zhang, Peiwen, et al. “A Study on the Utility Measurements and Influencing Factors of High‐speed Rail and Air Passenger Travel - Zhang - 2022 - Journal of Advanced Transportation - Wiley Online Library.” A Study on the Utility Measurements and Influencing Factors of High-Speed Rail and Air Passenger Travel, 16 Mar. 2022, onlinelibrary.wiley.com/doi/10.1155/2022/3224485. Accessed 13 Mar. 2025

Zhang, Qi, et al. “The Hub Competition in Delivering Air Connectivity between China and Oceania.” MDPI, Multidisciplinary Digital Publishing Institute, 25 May 2022, www.mdpi.com/2071-1050/14/11/6482. Accessed 13 Mar. 2025.

Nonprofits and For-profit Firms: Exploring Meaningful Work and Employee Fulfillment

By Claire Thrasher

Abstract

This research paper explores the relationship between employee work type (nonprofit or for-profit) and the perception of meaningful work, aiming to identify who feels the most fulfilled. Previous research shows that factors affecting employee satisfaction in nonprofit and for-profit firms have not been identified. I examined differences between for-profit and nonprofit firms using a cross-sectional survey analyzing career satisfaction, environment, stress, pay, situations, and perceived societal impact. While hypotheses anticipate higher meaningful work in nonprofit firms and acknowledge stress's impact, this research challenges prior assumptions about factors influencing meaningful work. The study concludes by emphasizing the need for further exploration through a larger sample size and the critical role of nurturing purpose-driven cultures in all workforces.

Keywords

Behavioral and Social Sciences, Psychology, Nonprofit, For-Profit, Meaningful Work.

Introduction

Many employees desire work that has a sense of purpose.[1] In this sense, they aspire for careers beyond earning “just money” and allowing them to impact the world around them. Hence, the term “employee satisfaction” is often used to express the level of fulfillment that employees experience regarding various aspects of their occupation - monetary and otherwise. These include the purpose of their current job, the overall holistic employee experience, and the organizations with which they are associated.[2]

One way that employees in organizations might achieve prosocial aims and feel more fulfilled in the first place is by working in more prosocial organizations, i.e., “nonprofits.” Nonprofit organizations play a significant role in fostering communities and strengthening society. They are typically driven by a mission or a cause and aim to serve the public or a specific community. Nonprofits also vary in size and form, in the sense that volunteers can run them with no financial resources or by skilled employees at large multi-million-dollar institutions.[3,4]

On the other hand, for-profit organizations often have differing goals from nonprofits - the ostensible goal is to maximize return on investment. Like nonprofits, however, for-profit businesses exist worldwide in various industries and range from small local businesses to major corporations. Employees play an important role in these kinds of firms as they contribute to the profitability and growth of these businesses. Additionally, employees often find meaningful or fulfilling work even in for-profit organizations.[5]

Claire Thrasher[1*], Jared Scruggs[2]

1. Carondelet High School, 1133 Winton Dr, Concord, CA, 94518, USA
2. University of Pennsylvania (UPenn), Philadelphia, PA, 19104, USA

Despite certain similarities in nonprofit and for-profit firms, a question remains - which employees feel the most fulfilled in their work, and why? Who is excited to go to their job every day? Research has not identified the factors that make employees feel the way they do about work. It is important to consider that many variables could affect an employee's satisfaction with their work. This includes their work environment, stress levels, home life, pay, etc.[1] Similarly, does the nature of the work itself - just being in a nonprofit vs. for-profit environment make a difference?

This paper explores the connection between employee work type and meaningful work. Most importantly, it aims to answer the question: Who feels the most fulfilled, nonprofit or for-profit employees? I do this using detailed questions throughout the survey. The cross-sectional survey analyzes employees' career satisfaction, situations, and perceived societal impact. By comparing participants' responses, a bigger idea will emerge regarding employees' fulfillment sources.

This study makes a major contribution. It aims to resolve a key question about employee meaningfulness. We might assume that a higher salary would cause employees to find more meaning in their work. This could also predict that individuals with less well-paid roles struggle to find meaning in their work. Nevertheless, it is important to recognize that meaningful work can come from more than just pay.[6] This research will offer insight into answering the main question: Who feels the most fulfilled, nonprofit or for-profit employees?

Background and Theory

Nonprofit and for-profit organizations are different organizational forms that operate with unique objectives and differing organizational structures. The main contrast between for-profit and nonprofit firms is their primary objectives, which commonly revolve around profitability - in the case of for-profit firms - or social impact for nonprofits. This difference significantly influences the strategies these firms adopt and even the employees they hire.[7] Although hybrid organizations exist that combine aspects of both, most organizations adhere to the previously outlined objectives for nonprofit or for-profit firms.[8]

For-profit firms play a significant role in the global economy and contribute to job opportunities. According to previous research, over 90% of all registered businesses worldwide are for-profit companies.[9] Globally, for-profit firms also employ around 70% of the workforce,[10] providing much-needed income to millions. Moreover, they can generate profits that benefit their corporations and facilitate long-term expansion.

On the other hand, nonprofit firms often address societal issues and advance the common good. In the U.S., nonprofits employ approximately 12.5 million people, accounting for about 10% of the American workforce.[11] The United States has over 1.5 million nonprofit organizations in many sectors. Their diverse focus areas encompass education, healthcare, environmental conservation, social services, and many other vital sectors, making them large contributors to the betterment of society.[12] Many employees often find meaningful work in these enterprises due to their focus on improving the world.

Meaningful Work: A Collective Journey

Within these organizations, both for-profit and nonprofit organizations, providing for or achieving meaningful work is an important endeavor for management and employees. Meaningful work is defined as work that is significant and positive in valence[13] and is often related to positive employee outcomes, such as higher self-esteem and improved overall well-being. Similarly, individuals feel a sense of purpose and fulfillment in their tasks - both strong correlates of meaningful work with high conceptual overlap and they are more likely to be engaged, motivated, and productive, leading to increased performance and overall organizational success.[14] This suggests that meaningful work is desirable, irrespective of work type.

Notwithstanding, due to for-profit vs. nonprofit firms' differing focus, key differences may exist in how employees experience meaningful work. Past research has found that profit-driven objectives in competitive industries can overshadow meaningful work. In for-profit organizations, the continuous effort to achieve financial gains often leads to the neglect of employee well-being.[15] Moreover, even in nonprofit organizations, the pressure to secure funding and maintain operations can occasionally weaken the focus on the meaningful impact of these organizations' missions. This can potentially jeopardize employee satisfaction and fulfillment, causing the sense of meaningful work that employees seek to disappear.[16] As such, balancing the demands of a competitive environment with the need for meaningful work remains a critical challenge.

Meaningful work goes beyond job responsibilities; it embodies the connection between employees' passions, values, and the impact they create in their professional careers. When employees are engaged in work that aligns with their sense of purpose, it elevates their overall well-being and fosters a sense of fulfillment. Organizations that foster purpose-driven cultures and empower employees will often benefit in terms of productivity and overall success.[17] Striving for meaningful work is a personal journey and a collective responsibility that enriches employees' lives and contributes to a more purposeful and thriving society.[15]

I hypothesize the following:
H1: The sense of meaningful work will be higher in non-profit than for-profit firms.
H2: Stress will have a significant impact on meaningful work.

Work Type Determines Meaningful Work

Despite the potential challenges of meaningful work in non-profit firms, non-profit firms have a large focus on meaning in general. Non-profit firms are likely to be higher in prosocially motivated employees and higher in prosocial contact with beneficiaries.[18] This priority of purpose in non-profit firms often attracts employees who are passionate and driven by prosocial aims. This results in an environment where meaningful work thrives and is encouraged in operations.[19] Generally, I anticipate this higher emphasis on meaningful work results in higher prosociality throughout non-profit firms.

However, resources are necessary for any job, and work without adequate resources can quickly become strained. The conceptual frameworks of the "job demands-resources model" and the comprehensive "self-determination theory" encompass the pillars of autonomy, competence, and relatedness and jointly emphasize the pivotal significance of resources in nurturing meaningful work.[20,21] Overall, when employees recognize low resources it is difficult to feel that their work is meaningful. This suggests that high resources can result in more meaningful work and a sense of fulfillment. This dynamic further displays a more pronounced illustrative demonstration of the distinctions between the nonprofit and for-profit sectors. The extent of this contrast will grow more significant as nonprofit workers' perception of possessing ample resources becomes heightened.[22]

Methods

Sample and Procedure

I surveyed 116 participants across various industries using convenience sampling on SurveySwap. Participants' median age was 24, and they worked an average of 30 hours per week. Their gender was 41 male, 73 female, and 2 did not disclose. Median household income was $20,000 - 29,999 (measured as an ordinal variable). Participants worked in a variety of types of jobs, but specifically, 67 participants were in private for-profit, while 27 were in either government (21) or private non-profit (6), 13 were self-employed, and 9 in other (e.g., "working without pay in the family business," etc.).

Measures.

Unless otherwise mentioned, all established scales were measured from 1 = "strongly disagree" to 7 = "strongly agree."

Independent variables.

For work type, I asked participants which kind of job they held, e.g. "private not-for-profit," "for-profit," etc. The complete list of options is available in the appendix.

Dependent variables.

For meaningful work, I used five items: four from Steger and colleagues (2012) positive meaning scale from the Work as Meaning Inventory (WAMI), e.g., "I have found a meaningful career," and one item which I added - "I feel my work is meaningful."
I also added three new items: "I enjoy working with my colleagues," "My work is challenging," and "I am interested in moving from my present job."

I used four items from Edwards and Rothbard (1999) regarding job satisfaction, e.g., "In general, I am satisfied with my job."

I adapted items from The Perceived Occupational Stress Scale (2021) for work and home stress. An exploratory factor analysis revealed that overall stress items are divided into these two

scales (work and home). Thus, I averaged four items for work stress, e.g., "My work is stressful," "My work has negative effects on my health," "Thinking about my work makes me feel tense," "At work, I feel under pressure," and three items for home stress, e.g., "my home life is stressful," "thinking about my home life makes me feel tense," "the stress that comes from my home life impacts how I feel at work."

Controls.

I also asked participants to report their income, gender, age, and hours worked per week to use in control models and robustness checks.

Results

Table 1 lists the main study variables' means, standard deviations, and correlations. The first analysis tested the sample's group differences between for-profit and non-profit groups. Although there were no significant group differences, a trend toward lower distribution (M = 3.62) and lower interpersonal justice (M = 5.18) for non-profit was reported as compared to for-profit (M = 3.93 and M = 5.43, respectively). Notably, however, these results only trended (not significantly).

Having established no significant differences between these groups, I wanted to understand if the relationships between justice and meaningful work differed within these samples. Income had a small but consistent association with meaningful work, $r(114) = .22$, $p = .017$, and this relationship was the same strength within for-profit or non-profit subsamples.

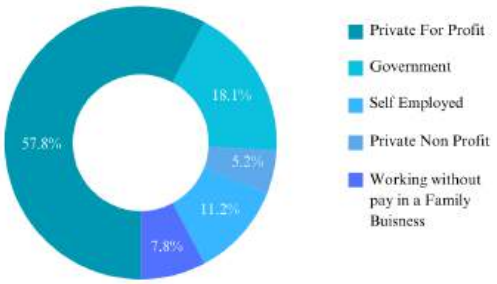


Figure 1. The number of survey participants as percentages by type of workforce. Most participants work in the private for-profit sector (57.8%).

Across both for-profit and non-profit, the relationship between distributive justice and meaningful work was moderate and positive, $r(114) = .36$, $p < .001$, as was the relationship between interpersonal justice and meaningful work, $r(114) = .31$, $p < .001$. However, while the relationship between interpersonal justice and meaningful work was the same within each subsample (for- or non-profit), within the non-profit subsample, the relationship between

distributive justice and meaningful work was not significant, $r(25) = .18$, $p > .1$, suggesting that employees in non-profit environments. At the same time, they may be sensitive (or not) to issues around pay and may not be affected by pay related to whether or not they find their work meaningful. Indeed, testing this as an interaction with distributive justice by work type (profit/non-profit), the interaction coefficient suggested a trending effect ($p = .2$) and the same simple slopes (a relationship between distributive justice and meaningful work for for-profit, and no relationship for non-profit).

Table 1. Study means, standard deviations, and correlations.

Variable	*M*	*SD*	1	2	3	4	5	6
1. Distributive Justice	3.89	1.61	--					
2. Interpersonal Justice	5.27	1.26	.34**	--				
3. Meaningful Work	4.64	1.38	.35**	.31**	--			
4. Job Satisfaction	4.69	1.30	.53**	.55**	.57**	--		
5. Work Stress	3.75	1.42	-.14	-.22*	-.01	-.26**	--	
6. Home Stress	3.48	1.52	.01	-.08	-.02	.11	.42**	--

N = 116.
*$p<.05$
**$p<.01$

Job satisfaction is positively correlated with distributive justice, interpersonal justice, and meaningful work. Work stress is negatively correlated with interpersonal justice and job satisfaction. Home stress correlates with work stress.

Discussion

Although there were no significant group differences between for-profit and nonprofit groups, some trends were uncovered. Within nonprofit firms, the relationship between distributive justice and meaningful work was as insignificant as in for-profit firms. This suggests that feelings about pay in nonprofits may not be as influential as predicted (Table 1). The evidence shows that the influence of income on meaningful work was minor, showcasing a significant contrast from the initially anticipated result.

Detailed Trends and Differences

While the distributive and interpersonal justice trends suggest differences between for-profit and nonprofit groups, caution should be exercised when interpreting these results. For example, there was a low sample size, which suggests there was low power to detect a significant difference. In addition, non-optimal combinations were made (e.g., government and private non-profit) to increase the size of the non-profit group for a higher power. However, it is intriguing to observe that the results differed significantly from prior expectations regarding the impact of stress and pay on meaningful work. More specifically, there was zero relationship

between stress and meaningful work in both groups, $p > .1.$, which contradicts previous expectations (Table 1).

In simpler terms, the research challenged prior assumptions about what makes employees feel fulfilled at work. The unexpected findings that stress and income have a minimal impact on meaningful work highlight a need for further exploration in this field (Table 1). The study prompts a need for a deeper understanding of the complexity of what makes work fulfilling for humans.

Some limitations in this study include all the information being measured simultaneously and measuring with a small sample size. Future studies should use larger sample sizes to draw definitive conclusions about the differences between nonprofit and for-profit organizations. They should also consider the proper delineation of groups and which groups are best combined or left separate. Still, implications like the lack of a relationship between stress and meaningful work and income having a small impact on meaningful work bring value to this study. Both for-profit and nonprofit sectors must foster purpose-driven cultures to empower and engage their workforce (Table 1).

Conclusion

In conclusion, this research challenges conventional assumptions regarding the factors that impact employee fulfillment in nonprofit and for-profit settings. Unforeseen results indicate limited connections between stress, income, and the perception of meaningful work. While some trends address distinctions in justice perceptions, the lack of statistical data emphasizes how much more research can be done in this field. Further investigation and exploration could prompt a more conclusive understanding of what fulfills humans in the workforce. This study evokes new research and represents the crucial role of nurturing purpose-driven cultures in all types of workforces (Table 1).

Acknowledgments

I would like to thank Polygence Research Academy for your support and guidance.

Works Cited

1. Steger, M. F., Dik, B. J., & Duffy, R. D. (2012). Measuring meaningful work: The work and meaning inventory (WAMI). Journal of Career Assessment, 20(3), 322-337.
2. Benz, M. (2005). Not for the profit, but for the satisfaction?–Evidence on worker well‐being in non‐profit firms. Kyklos, 58(2), 155-176.
3. Liebschutz, S. F. (1992). Coping by nonprofit organizations during the Reagan years. Nonprofit Management and Leadership, 2(4), 363-380.
4. Kong, E. (2008). The development of strategic management in the non‐profit context: Intellectual capital in social service non‐profit organizations. International Journal of Management Reviews, 10(3), 281-299
5. George, G., Haas, M. R., McGahan, A. M., Schillebeeckx, S. J., & Tracey, P. (2023). Purpose in the for-profit firm: A review and framework for management research. Journal of Management, 49(6), 1841-1869.
6. Dobre, O. I. (2013). Employee motivation and organizational performance. Review of applied socio-economic research, 5(1).
7. Rojas, R. R. (2000). A review of models for measuring organizational effectiveness among for‐profit and nonprofit organizations. Nonprofit Management and Leadership, 11(1), 97-104.
8. Ebrahim, A., Battilana, J., & Mair, J. (2014). The governance of social enterprises: Mission drift and accountability challenges in hybrid organizations. Research in organizational behavior, 34, 81-100.
9. World Bank. (2012). World Development Indicators 2012. The World Bank.
10. Foster, J. B., & McChesney, R. W. (2017). The global reserve army of labor and the new imperialism. Class: The Anthology, 485-502.
11. Cheese, P., Thomas, R. J., & Craig, E. (2007). The talent powered organization: Strategies for globalization, talent management and high performance. Kogan Page Publishers.
12. Faulk, L., Kim, M., Derrick-Mills, T., Boris, E., Tomasko, L., Hakizimana, N., & Nath, L. (2021). Nonprofit trends and impacts 2021. The Urban Institute.
13. Rosso, B. D., Dekas, K. H., & Wrzesniewski, A. (2010). On the meaning of work: A theoretical integration and review. Research in organizational behavior, 30, 91-127.
14. Satuf, C., Monteiro, S., Pereira, H., Esgalhado, G., Marina Afonso, R., & Loureiro, M. (2018). The protective effect of job satisfaction on health, happiness, well-being and self-esteem. International journal of occupational safety and ergonomics, 24(2), 181-189.
15. Veltman, A. (2016). Meaningful work. Oxford University Press.
16. Chetkovich, C., & Frumkin, P. (2003). Balancing margin and mission: Nonprofit competition in charitable versus fee-based programs. Administration & Society, 35(5), 564-596.
17. Ho, V. T., Wong, S. S., & Lee, C. H. (2011). A tale of passion: Linking job passion and cognitive engagement to employee work performance. Journal of Management Studies, 48(1), 26-47.

18. Grant, A. M. (2012). Leading with meaning: Beneficiary contact, prosocial impact, and the performance effects of transformational leadership. Academy of Management Journal, 55(2), 458-476.
19. Mirzaei, A., Webster, C. M., & Siuki, H. (2021). Exploring brand purpose dimensions for non-profit organizations. Journal of Brand Management, 28, 186-198.
20. Bakker, A. B., & Demerouti, E. (2007). The job demands‐resources model: State of the art. Journal of managerial psychology, 22(3), 309-328. https://doi.org/10.1108/02683940710733115.
21. Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. Handbook of theories of social psychology, 1(20), 416-436.
22. Fairlie, P. (2011). Meaningful work, employee engagement, and other key employee outcomes: Implications for human resource development. Advances in developing human resources, 13(4), 508-525.

The Impact of Single-Panel Storytelling on Human Culture: 50,000 BCE to the AI Age
By Alex Kim

Abstract

Predating written history, visual storytelling functions as one of the central pillars upholding human culture through communication. Its evolution through the millennia of prehistory has taken humans through technological and cultural epochs, from the oldest discovered cave paintings to the AI-generated art now permeating the internet. Single-panel art specializes in illustrating moments, with the power to convey knowledge, express emotion, and convey tradition. This paper surveys the anthropological and sociological roles of single-panel art as a storytelling medium, emphasizing its significance in creating cultural artifacts in any time period. The scope of this article includes the constancy of single-panel art's role as well as its transformations in the digital age. Emerging digital tools and artificial intelligence must be understood to continue serving artists and their audiences as storytelling democratizes with increasing channels for audience engagement. The ongoing impact of single-panel art in visual storytelling as a narrative vehicle necessitates further discussion of complex factors that both threaten and provide opportunities for artists worldwide. In conclusion, this paper introduces the ethical issues surrounding the current ecosystem of art in relation to AI, underscoring the potential value of digitization as a means of both preservation and expression as long as traditional artistry remains protected.

Introduction

The concept of storytelling is built into human history, proven by the existence of the time period known simply as pre-history—the time before written languages recorded human events. Today, the concept of telling a story falls under countless genres including drama, tragedy, comedy, and mystery, and the forms that let these stories be told are just as numerous. Humans express their creativity, spread knowledge, pass down traditions, and establish cultural values through the stories they tell through oral, written, and visual representations of their thoughts. Of all the forms, something striking about the visual arts is its non-reliance on technology and its usefulness in avoiding "the problem of storage capacity that orality faced" when recording and increasing collective knowledge (Armstrong 2). Innovations throughout time have only increased the tools available to artists. Visual art fascinates the eye, as demonstrated by the oldest discovered cave paintings, fine artworks on the canvas, and digital illustrations like memes, emojis, and AI-generated art.

Single-panel art, or images completed within one frame that can stand alone, has evolved with human communication and culture. From depictions of successful hunts to creative expressions of a mind's inner workings, illustrated moments remain a valuable tool in describing our surroundings and defining our world, giving useful information like technical knowledge or perpetuating culture and traditional norms (Hurlburt and Voas 4). Images also reflect the simple in the complex, with a single frame of art containing a world of thoughts and feelings in itself, as

captured in a stunning Olympic photo finish or the hardship of a mother during the Great Depression. Within one image, an entire story can be told.

Technological advancements have ushered in a revolution of single-panel storytelling, with the spread of digital art tools and online circulation making art more interactive and democratic than ever before: audiences can engage with art and artists through online platforms on a scope never seen before the late-20th century (Anderson and Chua 32). Traditional art and the birth of digital single-panel media, such as webcomics, social media memes, and AI-generated art demonstrate the ongoing significance of visual storytelling to this day. Hence, this paper examines the cultural roles of single-panel art, the emergence of digital technology, and the changing ecosystem of art moving into the age of artificial intelligence.

The Role of Single-Panel Art in Storytelling

Visual narratives amplify the human reaction to imagine, interpret, and create. Single-panel art plays a crucial role in communication because art transcends all languages. From cave paintings to iconic images of the divine, to political cartoons and photographs, lone images are powerful vehicles carrying informative and persuasive power; the shock or pathos of an image can altogether overpower logical strings of rhetoric. Though a single image may not tell the complete story, it captures the essence of what an artist needs to tell, inspiring further engagement. This section details the anthropological and sociological dimensions behind what makes single-panel art stand out throughout our shared human history.

Anthropological Drive

Many people love to talk, and others listen. What all people have in common is that they have a story to tell, a story that is influenced by other stories that exist before or will be told after theirs. Evidence throughout archaeological studies shows that the drive to tell stories has played a significant part in humanity's rise as a species (Vinall). Predating written history, cave paintings depicting more than hunting victories have been recently discovered and dated as nearly 52,000 years old in Indonesia and, according to Will Sullivan of the Smithsonian, "may be the earliest known example of visual storytelling." Their existence validates the claim that human civilization centers around the exchange of information beyond verbal forms (Ayub et al. 68).

As archaeologists have discovered and anthropologists have learned, humans have told stories throughout the existence of groups. Despite the condition of a written record to be counted as history by historians, anthropologists confirm prehistoric oral tradition through the remnants of pictorial history: the accompaniment of a told story by illustrations that record notable moments of forgotten "primitive and peasant cultures [who] were thought to be without history" (Ulin 390). However, anthropology now demonstrates the important interactions of history's genuine storytellers who through gestures and illustrations succeed in meeting "social, cultural, physical, and/or emotional needs" by providing their audiences with crucial context (Ryan 69). Thus, the anthropological drive behind storytelling is actually a two-part

participation, as storytellers and audiences both learn, engage, and share in a unique cultural exchange. The addition of visual elements allows more people to understand, interpret, and imagine more closely in line with the intentions of a story's truth.

Throughout recorded history, especially with the development of fine art during the late Middle Ages (1200-1500 CE), single-panel artistry grew with global exchange to enrich storytelling as Renaissance artists emerged. The motivation to reveal beauty and recognize ancient wisdom enabled the artists and sculptors to romanticize history as a story featuring great individuals. Iconic single works stand as testament to the power of a single image in both "documenting historical events and representing cultural values" as material artifacts ("How Sculptures Have Defined Moments in History"). For the sake of inclusion in this paper, sculpture is treated as three-dimensional single-panel art.

Even as media continue to evolve into the digital age, the value and use of single-panel art increases beyond traditional art and sculpture. With the proliferation of internet usage and digital tools, the modern genre of digital storytelling blends narrative creation with digital and internet-derived image use, creating an effective and easy method for people to represent cultural phenomena, ultimating "providing a description of the knowledge and practice of the group of people exposed to it" (Thornburg 182). On the global scale of the online world, this era empowers individuals who do not consider themselves artists to gain abilities to present their own narratives through low-barrier techniques like memes, screenshots, emojis, and art created through artificial intelligence. It is certain that the anthropological fascination with single-panel visual storytelling is flourishing.

Sociological Value

Because of its value in adding to the collection of global material culture, single-panel art and storytelling cannot be labeled as just an artistic skill. More than ever before, it is a symbolic interaction between storytellers and their audiences, a preferred medium to transmit thoughts and ideas (Polletta et al. 112). Moreover, images can present powerful messages to impact society, especially in digestible forms like comics that tackle real-world issues (Al Faruque 5). For example, single-panel art has served as social commentary for generations, as seen in Gary Larson's The Far Side cartoon series, political critique in nationalized editorials featuring political cartoons, and campaigns to highlight societal needs like Benjamin Franklin's colonial "Join, or Die" depiction of the American colonies.

Furthermore, single-panel art goes beyond documenting the times because cultural artifacts contain historical meaning, but their significance and interpretation evolve as new societies evaluate them. The meaning contained in a single visual image creates a shared story for every group interacting with it. According to Pescarin et al., the common cultural heritage stored within an image fosters greater social cohesion, meaning that specific images may bond people of a subculture together more intimately, just as universally understood images bring different large groups of people closer together (1). Single-panel art features heavily in religious iconography, propaganda, and social campaign flyers, demonstrating power in sharing narratives

to bond people to a purpose and vision, as well as mobilize action (Polletta et al. 118). In education and entertainment, a single image accompanying literature can amplify the impact of a plot's climax or exemplify a moral lesson in a cautionary tale, with the added benefit of sparking the imagination. Just as important, visual stories ensure the idea of the content remains even after the words of a story are forgotten or suppressed, inspiring the imagination to fill in when historical truth is unclear (Hurlburt and Voas 4).

Finally, in the modern world, where data bombard the ears and eyes, single-panel visual storytelling excels at both immediate and intimate information exchange. Single images are meant to deliver messages instantaneously, and they succeed even further when they inspire lingering thoughts that coax greater recognition and emotional impact. In the digital age where creators and commenters not only coexist but also often interact, connections are actively made, and such lingering thoughts are expressed instead of remaining unvoiced by passive audiences. Therefore, digital visual storytelling accelerates the circulation of shared narratives and makes storytelling a more collective process.

Technology and Digitization

The evolution of single-panel art from cave paintings to digital AI art exemplifies a common purpose in every time period, which is to preserve and express information. The modern preservation of visual art began with the digitization of visual information in the 1980s, but the following decades have ushered in the use of digital tools to create art and democratize original and artificial content creation.

Means of Preservation

When speaking of art, preservation carries two meanings. First, art conveys and preserves information in a form different from oral and written communication. Second, art pieces themselves are to be preserved to remain meaningful cultural artifacts that can still convey information. As archaeologists and anthropologists know all too well, traditional preservation of art suffers the burden of working against the damages of time. Whether cave paintings survived was a matter of erosion and weathering, and the restoration of Renaissance paintings involves its own debate on how much original work remains after repair. Thus, digitization of traditional art pieces and their ease of replication serves as an active kind of knowledge preservation (Ayub et al. 73). Whenever art is photographed, scanned, and turned into a document or file, it is preserved as a copy, and even its own copies can be duplicated as reliable identicals.

AI: Democratization or Transgression?

Beyond taking screenshots to preserve snippets of information or circulate captured moments, technology is being used to create art digitally, with recent advancement allowing the automation and generation of content through artificial intelligence.

Single-panel art has become more accessible both to view and to emulate ever since the rise of the computer screenshot. A standard feature on computers, mobile devices, and

smartphones, the screenshot function and pre-installed editing software have evolved into a new form of user-generated single-panel art and cultural artifact, with purposes including documentation, communication, and self-expression (Mottelson 741). The increased accessibility of this technology has shifted how art is created and expressed: though socially exchangeable art once used to be based on its technical merit (is the art good enough to share?), user-generated digital panels are more personally evaluated (is this screen capture worth sharing?). The growing participation in creating and circulating art has been called a new democratization of art as it is becoming more possible for anybody to create, modify, and share art regardless of ability.

The newest development in digital creation is generative art, or AI art. Artificial intelligence learning models—often trained with extensive access to traditional art images—have been leveraged to create art based on verbal prompts, resulting in what AI advocates call the most complete form of art democratization. However, traditional artists criticize AI developers in the belief that generative AI represents artificial creativity through its synthesis of multiple styles to create a homogenized look that muddles artistic genres. In his paper discussing the challenges and opportunities presented by generative AI in the art industry, Manuel Garcia notes the importance of finding the agreeable overlap where traditional art perspectives and generative artificial intelligence can synthesize to truly democratize art through true collaboration that boosts the economic prospects of artists, respects their creator's rights, and recognizes accepted ethical boundaries so that artists, subjects, and audiences are represented equitably (5, 17). While the ongoing debate is fierce, it is still rather new, deserving ongoing examination in the interests of protecting traditional artists and making a way for the technology to prove useful and non-threatening.

Conclusion

From cave paintings to AI-generated art, the purpose of single-panel art is clear: to tell and preserve stories. As we want to believe, social media is a unique technological innovation, but it largely serves the same purpose as the earliest artists etching singular images into rock walls. Compared to what the ancients did, today's people engage in the same activities, just in a modern form that embraces the tools of our time. In the same way primitive peoples shared their achievements through creative expression, we too share our chosen moments through curated photographs on a social media timeline and express ourselves with the perfect emoji. Those who lack the skills to portray their thoughts by brush have even taken to AI to make their ideas come to life. Thus, despite the potential threats against traditional art in the age of AI, the opportunities to democratize art and increase the potential for visual storytelling are very real, if used ethically and with respect to creator's rights. In any era and any form, the power of single-panel art in telling stories remains undeniable, with the societal benefits of preserving culture and encouraging communication.

Works Cited

AARNE. “How Sculptures Have Defined Moments in History.” AARNE Bespoke Sculptures, 12 July 2023, https://aarne.co.uk/how-sculptures-have-defined-moments-in-history/.

Al Faruque, Saidya Nasira. “The Unique Power of Comics: A Comprehensive Exploration of the Visual-Verbal Medium and Its Impact on Storytelling, Communication, and Culture.” International Journal of English and Studies, vol. 5, no. 10, 2023, pp. 1-11. International Journal Of English and Studies (IJOES), DOI:10.47311/IJOES.2023.10.11.

Anderson, Kate T., and Puay Hoe Chua. “Digital Storytelling as an Interactive Digital Media Context.” Educational Technology, vol. 50, no. 5, 2010, pp. 32-36. JSTOR, http://www.jstor.org/stable/44429857.

Armstrong, Grace. “Cave Art and the Development of Written and Visual Communication.” KINO: The Western Undergraduate Journal of Film Studies, vol. 8, no. 1, 2019, pp. 1-9. Western Undergraduate Film Society, University of Western Ontario, https://ojs.lib.uwo.ca/index.php/kino/article/view/13618.

Ayub, Aamir, et al. “Evolution of Communication: A Comparative Study from Cave Paintings to Emojies.” Qlantic Journal of Social Sciences, vol. 5, no. 3, 2024, pp. 68-74. Qlantic Journal of Social Sciences, https://doi.org/10.55737/qjss.770639507.

Garcia, Manuel B. “The Paradox of Artificial Creativity: Challenges and Opportunities of Generative AI Artistry.” Creativity Research Journal, 2024, pp. 1-14. Taylor & Francis Online, https://doi.org/10.1080/10400419.2024.2354622.

Hurlburt, George F., and Jeffrey Voas. “Storytelling: From Cave Art to Digital Media.” IT Professional, vol. 13, no. 5, 2011, pp. 4-7. IEEE Xplore, DOI: 10.1109/MITP.2011.87.

Mottelson, Aske. “Why Do People Take Screenshots on Their Smartphones?” DIS '23: Proceedings of the 2023 ACM Designing Interactive Systems Conference, 2023, pp. 740-52. ACM Digital Library, https://doi.org/10.1145/3563657.3596067.

Pescarin, Sofia, et al. “Social Cohesion in Interactive Digital Heritage Experiences.” Multimodal Technologies and Interaction, vol. 7, no. 6, 2023, pp. 1-20. Multidisciplinary Digital Publishing Institute, https://doi.org/10.3390/mti7060061.

Polletta, Francesca, et al. “The Sociology of Storytelling.” Annual Review of Sociology, vol. 37, 2011, pp. 109-30. JSTOR, http://www.jstor.org/stable/41288601.

Ryan, Patrick. “The Storyteller in Context: Storyteller Identity and Storytelling Experience.” Storytelling, Self, Society, vol. 4, no. 2, 2008, pp. 64-87. JSTOR, http://www.jstor.org/stable/41949003.

Sullivan, Will. “Indonesian Cave Painting Is Oldest-Known Visual Storytelling.” Smithsonian Magazine, 5 July 2024, https://www.smithsonianmag.com/smart-news/indonesian-cave-painting-is-oldest-known-visual-storytelling-180984660/.

Thornburg, Aaron. “Digital Storytelling as Autoethnography in Anthropological Pedagogy and Practice.” Deep Stories: Practicing, Teaching, and Learning Anthropology with Digital

Storytelling, edited by Aaron Thornburg, et al., 1 ed., De Gruyter, 2017, pp. 179-90. JSTOR, http://www.jstor.org/stable/j.ctvbkjvdr.16.

Ulin, Robert C. “The Anthropologist and the Historian as Storytellers.” Dialectical Anthropology, vol. 19, no. 4, 1994, pp. 389-400. JSTOR, JSTOR, http://www.jstor.org/stable/29790380.

Vinall, Frances. “51,000-year-old Cave Painting May Be Earliest Scene Depicted Through Art.” The Washington Post, 3 July 2024, https://www.washingtonpost.com/science/2024/07/03/oldest-cave-painting-scene-indonesia/.

The Dual Transformation:
How Technology and Sustainability are Reshaping the Fashion Industry By Riddhi Jain

Research Question

To what extent are sustainability considerations and technological advancements reshaping practices within the fashion industry?

Abstract

This paper explores the extent to which sustainability considerations and technological advancements are reshaping practices within the fashion industry. It examines the industry's historical challenges with sustainability, such as unethical labor practices, overproduction, and environmental degradation, alongside the rising influence of consumer awareness and global initiatives like SDG 12. The study highlights how brands such as Patagonia, Stella McCartney, and Eileen Fisher are adopting sustainable innovations, including recycled materials and circular initiatives. Simultaneously, Industry 4.0 technologies like AI, blockchain, and AR/VR are revolutionizing production and consumer engagement, exemplified by campaigns like Jacquemus' AI-driven digital creativity and Metaverse Fashion Week. These advancements underscore the inseparability of technology and sustainability in shaping fashion's future. By adopting these dual transformations, brands are not just responding to demands for responsibility but redefining the industry's standards for innovation and progress.

Key Words

Fashion industry, sustainability, technological advancements, Industry 4.0, responsible consumption.

Introduction

With the need to balance heritage, innovation, and sustainability, is fashion at a crossroads?

The fashion industry is a dynamic fusion of culture and creativity, reflecting societal values and global trends through innovation. The industry originated in the 19th century with Charles Frederick opening the first haute couture house in Paris, and entered a new era as a result of the Industrial Revolution, which introduced mechanised textile manufacturing, allowing quick production and distribution of mass-produced garments, causing a rise in fast fashion in the late 20th century (Harper's Bazaar, 2017). Today, this industry is valued at $1.84 trillion and continues to grow, employing over 70 million workers (Licata and Kemp, 2020). However, the fashion industry is anything but static and straightforward - its global reach necessitates evolution, given its exposure to varying macroeconomic changes.

In recent times, there has been an increasing awareness of fashion's environmental and social impact, given the statistics surrounding the industry's depletion of resources, excess pollution emissions, and unethical labour practices. For instance, it has been reported that the

industry produces 92 million tonnes of textile waste each year (Igini, 2023). As a result, fashion companies are now compelled to integrate sustainable business models, including circular economy principles, to mitigate their environmental footprint by improving energy management, deploying new infrastructure, and practising fair trade (Roy and Iacoviello, 2024). Simultaneously, another macroeconomic factor changing the traditional operations of the fashion industry is technology. Artificial intelligence (AI), Augmented Reality (AR), Virtual Reality (VR), and blockchain, amongst other technologies, are revolutionising the design, production, and customer engagement of fashion brands (UNSW, 2023). This paper, therefore, aims to answer the following question: **To what extent are sustainability considerations and technological advancements reshaping practices within the fashion industry?**

This paper, using real-life examples, will argue that the fashion industry must adopt sustainable practices and integrate technological innovations across its supply chain to remain competitive, meet changing consumer demands, and align with global sustainability goals.

Literature Review: The Macroeconomy and PESTLE Analysis

Internal factors, such as organisational culture, management, operations, finance, and workforce, directly affect business performance. External factors, on the other hand, are factors outside a business's control, also known as the macroeconomy. It includes factors such as inflation, GDP growth, employment rates, and monetary policies that shape industries at a national and global level (The Investopedia Team, 2023). A key method of assessing macroeconomic factors is through the PESTLE framework, which represents Political, Economic, Social, Technological, Legal, and Environmental trends that impact a business's market share and position in the marketplace. This business management tool helps owners and project managers in the decision-making process as it helps in identifying industry trends, mitigating risk, leveraging growth opportunities, and developing adaptive strategies (Makos, 2024).

FACTOR	EXPLANATION
POLITICAL	Political stability and government activity are associated with factors such as the sentiment towards the chosen government, trade regulations, taxation, and the legal framework. While strong government support through flexible policies can support firms' growth, changes in political leadership can impact market conditions and investor confidence, and weak or corrupt government operations can reduce business efficiency.
ECONOMIC	The state of the economy is vital as it influences consumers' purchasing power, the market's suitability for businesses, and access to factors of production such as capital. For instance, an increase in economic growth would increase consumer spending on goods and services, improve logistics, and encourage investments. However, it may also cause inflation, increasing production costs and pricing instability, affecting profitability.

SOCIAL	Social factors shape business through demographics, cultural values, and consumer behaviour. A growing and educated population can increase the quantity and quality of the labour force as well as increase awareness among consumers. However, rapid societal changes can drive up costs, while increasing income inequality restricts the purchasing power of lower-income consumers. Shifts in cultural demographics may also spark controversies. Furthermore, unethical business practices can tarnish brand reputation and negatively impact financial performance.
TECHNOLOGICAL	Technological factors refer to advancements in innovation, digital transformation, and research and development. The development of new technology can influence business processes, supply chains, and product choices. However, it might also cause cybersecurity and privacy issues with data.
LEGAL	The legal framework regulates business operations, employment, and consumer protection. Effective trade and labour laws can increase business activity. Taxation, competition, and data protection regulations can allow for operational transparency and market stability. However, high taxes on goods and services can discourage consumer spending in the economy, reducing business confidence and lowering GDP.
ENVIRONMENTAL	Sustainable business practices can improve a business's long-term profitability and brand image. They can also encourage firms to use green energy practices, leading to cost savings and waste management. However, climate change and natural disasters can disrupt supply chains and halt production.

Expanding more closely on environmental factors, sustainability is a key word that has arisen. Sustainability refers to meeting the needs of the present without compromising the ability of future generations to meet their own needs by ensuring the long-term preservation of environmental, social, and economic resources (United Nations, 1987). Social sustainability focuses on equality, human rights, and community well-being. Economic sustainability ensures the business remains viable, generating jobs and economic stability without exploiting resources or workers. Economic and social sustainability are often not fully achieved due to companies having a short-term economic focus, a lack of government regulations, and low consumer awareness. Consequently, challenges such as high costs, labour exploitation, and greenwashing arise.

A framework like the triple bottom line guides companies to pursue sustainable development by balancing all three factors. The triple bottom line is a business concept that can be broken down into three Ps: profit, people, and the planet. This approach fosters ethical labour practices, resource efficiency, and long-term resilience, assisting businesses in designing operations that generate inclusive growth and ecological sustainability (Miller, 2020). This approach is vital for companies to employ as businesses face growing pressures from stakeholders. For instance, consumers are becoming more aware and prefer businesses that cater

to all sustainability dimensions. In line with this, research also shows that sustainability enhances brand value and customer loyalty. Moreover, the government and the legal landscape also necessitate adherence to sustainable practices, with rigid laws being put in place on emissions, reporting, and labour rights (Spiliakos, 2018). By integrating this core strategy, organisations can avoid risks and improve supply chain ethics and resource efficiency (United Nations, 2025).

Sustainable commitments by organizations are also imperative to meet the 17 Sustainable Development Goals (SDGS), a global initiative enforced by the United Nations to achieve a more sustainable future. Of these, SDG 12, which is based on responsible consumption and production, illustrates efficient use of resources. It focuses on transforming international consumption and production habits by promoting a circular economy, helping reduce waste, and increasing allocative and productive efficiency. It calls for producers and consumers to collaborate in making ethical and viable decisions (UNEP, 2021). The aforementioned has specific relevance in shaping companies' policies.

Moving beyond environmental factors, technological factors are also ever-changing in today's macroeconomy. Driven by disruptive trends, including the rise of data and connectivity, analytics, human-machine interaction, and improvements in robotics (McKinsey & Company, 2022), Industrial Revolution 4.0 especially marks a transformation in global businesses' operations. Key technologies emerging from Industry 4.0 include artificial intelligence (AI), automation and robotics, blockchain, and immersive technologies such as augmented reality (AR) and virtual reality (VR). AI enables smarter decision-making, predictive analytics, and personalised customer experiences while automation and robotics streamline operations, reduce labour costs, and significantly boost efficiency. In parallel, blockchain amplifies transparency and data security in the production process. Furthermore, through the help of AR and VR, 3D technologies are designed to stimulate real-world experiences. Businesses are revolutionising customer engagement, employee training, and product visualisation by offering immersive experiences that bridge the gap between digital and physical environments (UNSW, 2023). For instance, IKEA’s AR app allows customers to place virtual furniture in their homes using their smartphones or tablets and see the look of different furniture pieces in their space before making a purchase. This enhances the shopping experience by providing confidence and an interactive way of selecting products, consequently increasing engagement and sales. It also helps with driving up both customer satisfaction and customer loyalty rates (Gibson, 2024).

Collectively, these mechanisms - both environmental and technological - empower businesses to operate with greater agility and innovation. They help companies become more responsive to customer needs and preferences. By leveraging these tools, businesses can unlock new revenue streams and explore fresh business models (Gibson, 2024). Additionally, they contribute to improved sustainability through smarter resource use and reduced waste, reducing the average cost of production (Gibson, 2024). Therefore, amidst evolving macroeconomic challenges, embracing both environmental responsibility and technological innovation has become an imperative pillar for industries seeking resilience, relevance and long-term growth in the face of rapid changes.

The Impact of Sustainability on the Fashion Industry

The fashion industry has long been criticised for excessive dependence on rapid production and consumption cycles. Fast fashion, driven by low-cost and high turnover models, has led not only to the degradation of scarce resources but also to labour exploitation. At the heart of the fashion industry, unethical practice is the relentless drive for profit maximisation and the need to dominate the market share. To stay competitive in this industry, brands often compromise with cheap materials as well as outsource production to regions with minimal labour protection. This cost-cutting makes the company's products affordable, encouraging customers to purchase frequently and discard clothes after limited use. As a result, many fast fashion brands evade accountability, masking harmful practices behind vague sustainability claims. This can be seen in the case of brands like H&M and Shein, which have faced scrutiny that exemplifies this trend. H&M faced allegations of greenwashing following the launch of its "Conscious" collection, which claimed to use sustainable materials. Nonetheless, later investigations revealed that many of the collection's products contained a higher percentage of synthetic materials than their standard line, misleading consumers about their environmental impact (Pearcy, 2021). Meanwhile, Shein has been the subject of extensive media stories for labour practices that include excessive working hours, inadequate wage rates, and poor working conditions (Saunders, 2025). Reports indicate that workers in a few of Shein's factories work up to 75 hours per week and are often given only a day of leave per month, which is a violation of labour laws that suggest a maximum of 40 working hours a week. This leads to increasing backlash from consumers and environmental groups, forcing the companies to reevaluate their practices (Capella University, 2024).

However, rising consumer awareness is now driving a significant shift towards ethical and environmentally friendly practices. The fashion industry is undergoing a foundational transformation as sustainability is being integrated across the supply chain at every stage, from the raw material selection to the methods of production, distribution, and retailing. These changes are both driven by internal values and external pressures to meet regulatory standards and commit to ethical practice.

With regard to the sourcing of materials, transparency has become a cornerstone of sustainability - manufacturers are increasingly sourcing eco-friendly and ethically extracted materials. Brands now use blockchain and digital identities to trace garment origins, ensuring supplies adhere to fair labour practices. Moreover, traditionally, fashion relied on virgin cotton and synthetic fabrics like polyester, causing immense water consumption and the creation of microplastics. In a bid to be more sustainable with the materials being used, brands like Patagonia have pioneered the use of recycled materials like post-consumer plastic bottles, and have limited their environmental footprint in sourcing. Additionally, Adidas's ongoing partnership with Parley for the Oceans, which was launched in 2015, has produced over 30 million pairs of shoes made from plastic waste recovered from the ocean. This initiative not only mitigates ocean pollution but also raises awareness about the environmental impacts of unsustainable large-scale productions (Parley, 2024). Both companies have reduced reliance on

virgin resources, promoting sustainable management of natural resources and reducing pressure on ecosystem resources.

In the production phase, Stella McCartney not only prioritizes ethical practices, such as ensuring safe factory conditions and prohibiting child labor, but it is also the first luxury fashion brand to completely eliminate the use of animal leather, feathers, fur, or skin. Instead, the brand utilizes 100% cruelty-free materials, including plant-based and lab-grown alternatives. Committed to a closed-loop system, Stella McCartney incorporates waterless dyeing technologies and zero-waste patterns (Williment, 2025). Furthermore, the brand ensures the environmentally sound management of chemicals, minimizes hazardous waste, and actively prevents soil and water pollution. This reflects the principle of the triple bottom line - balancing profit with care for people and the planet.

At the retail level, consumers are directly being encouraged to participate in sustainability through take-back programs such as Eileen Fisher's Renew program, launched in 2009, which invites customers to return used garments in exchange for store credit. The returned items are then cleaned, repaired, and either sold or repurposed into new products, reducing the pressure on landfills (Eileen Fisher Renew, 2025) and making a significant contribution to textile waste reduction and promoting the 3 R's - Reduce, Reuse, Recycle. The brand further encourages consumers to shop second-hand through its renew platform, fostering a more sustainable wardrobe. This cultivates customer loyalty towards the brand, plus turns sustainability into a shared consumer practice.

Although the integration of sustainability may lead to high initial costs, it is no longer an option but a necessity for survival, to ensure long-term success and competitiveness as ethical concerns intensify. Brands that fail to adopt rational and sustainable practices in their supply chains risk losing customer trust, which can lead to public backlash, boycotts, and a competitive disadvantage compared to brands that prioritize sustainability. Moreover, the lack of sustainable practices may expose companies to increased government regulation, resulting in legal complications and operational setbacks. This negligence can also cause businesses to miss out on long-term growth opportunities, ultimately diminishing their profitability.

The Impact of Technology on the Fashion Industry

For decades, the fashion industry has thrived on in-person retail experience, seasonal physical runway shows, and tactile consumer engagement. However, this paradigm is rapidly shifting. The fashion industry, once rooted in tradition and manual craftsmanship, is now undergoing a rapid change. Emerging technologies such as AI, AR, VR, and blockchain are transforming how fashion is designed, produced, marketed, and consumed. While craftsmanship remains central, innovation is pushing the boundaries of digital engagement and virtual commerce, reshaping fashion's identity and its relationship with consumers. One of the most significant areas affected by this change is marketing.

In the past, fashion was primarily promoted through high-end magazines, billboard advertisements, and television commercials, where brands curated editorials of the latest

collections and fashion week features, reaching their target audience. With the rise of the internet and social media platforms, brands have adapted, moving their promotion strategies to online platforms - embracing platforms like Instagram, TikTok, and YouTube to engage with consumers directly, in real time, on a global scale. Social media, tailored for virality, emotional connection, and cultural relevance, allows consumers to no longer just observe fashion; they interact with it, remix it, and become a part of the narrative. Through the use of AI generative short visuals, algorithm-friendly videos posted in the form of Instagram reels, YouTube shorts, and TikTok allow brands to post behind-the-scenes content, tutorials, and teasers. Influencer collaborations have particularly become a cornerstone of digital marketing, with 64% of marketers working with micro influencers due to their high engagement rates and authentic connections with audiences (Wallace, 2025). Jacquemus exemplifies a fashion brand leveraging social media to its fullest potential. Recently, the brand collaborated with cultural icon Bad Bunny, aligning its image with modern masculinity and expanding its demographic reach (Hughes, 2023). Jacquemus has also pioneered innovative marketing campaigns that utilize digital technology and social media effectively (Tacticone, 2024). One standout campaign featured surreal, AI-generated 3D renderings of oversized versions of its iconic 'Bambino' bag rolling through the streets of Paris, as seen in the image below taken from (Mistry, 2023) - captivating audiences and amassing over 2 million views on TikTok (Sky Society, 2024). By creating cultural moments, Jacquemus achieves significant brand growth with minimal reliance on traditional advertising methods, boosting both its visibility and customer base.



Just as Jacquemus reimagines promotional strategies, the industry is now also rethinking the runway. Fashion weeks, once defined by exclusive front row seats and elaborate physical setups, are being reinterpreted for the digital age. Metaverse fashion week is a digital fashion event hosted on Decentraland, a blockchain-based virtual world offering an immersive platform where global audiences can attend virtual runway shows (as seen in the image below, on the left, taken from Guinebault, 2022), explore interactive pop-up stores (as seen in the image below, on the right, taken from Zangana, 2022), purchase NFT-based fashion drops, and enjoy avatar customisation, integrating AR and VR experiences for an enhanced sensory experience. The first season of this fashion week - hosted from March 24th to March 27th in 2022 - was deemed a massive success featuring over 70 brands including luxury names such as Dolce & Gabbana, Tommy Hilfiger and Etro (Thompson, 2023) and attracting approximately 108,000 attendees and facilitating the sales of over 7,000 virtual wearables, generating $76,757 (Klein, 2022). So, what led to its success? At its core, the Metaverse fashion week offers global accessibility, eliminating geographical and socioeconomic barriers long associated with traditional fashion weeks. Moreover, due to its virtual nature, there is a significant reduction in environmental footprint as emissions from travel, food, and accommodation are cut down. It also gives designers a platform to explore boundless creative expression without making prototypes, limiting the wastage of material resources (MacRae, 2025). Most importantly, the Metaverse fashion week introduces brands to a novel revenue model and deepens digital engagement, aligning with the upcoming lifestyle changes (Birch, 2021).





On the question of whether virtual fashion shows are going to replace traditional fashion shows entirely, however, this may not be any time soon, as there are still certain hurdles to overcome. Despite its open-access ethos, the meta fashion week remains niche, resonating mainly with gamers, crypto enthusiasts, and digital natives, excluding those with limited digital literacy or inadequate access to technology. Furthermore, although the technological aspect reduces the carbon footprint and wastage of resources, environmental concerns persist due to the energy-intensive nature of the blockchain operation (Dolan, 2022). To add to this, many consumers remain sceptical about paying for intangible fashion items, raising questions regarding the value and practical utility (Park, Ko, and Do, 2023).

Engaging with technology and utilizing it effectively are crucial for fashion brands, enabling them to achieve new heights in brand visibility and consumer engagement through virality and immersive experiences. As technology continues to evolve, it is likely to become more accessible, making its integration into operations even more essential.

Conclusion

This research paper aimed to explore the extent to which sustainability and technological innovation are reshaping industry practices. The consensus is that to stay viable and competitive, fast-fashion brands must embrace both sustainability and technology across their supply chain for long-term viability.

Fashion's historically unethical practices are marked by overproduction, exploitation of labour, and environmentally destructive processes. These practices have not only drawn public scrutiny but have also exposed brands to reputational risk. However, growing consumer awareness and global movements such as SDG 12, which emphasise responsible consumption and production, have catalysed conscious operations. Brands are now reforming; Patagonia uses recycled materials, Stella McCartney practises cruelty-free production methods, and brands like Eileen Fisher are introducing circular initiatives such as garment take-back programs. These changes respond not only to shifting consumer expectations but also to global sustainability requirements. It must be reiterated that these sustainability efforts are no longer an option; they are integral to a brand's legitimacy and market survival.

Parallel to these developments, Industry 4.0 technologies are revolutionising the industry from both a production and consumption standpoint. Through the help of advanced tools such as AI, blockchain, and 3D printing, raw material sourcing and manufacturing processes have become more efficient. Further, tools such as AR and VR have increased consumer engagement immensely. Jacques exemplifies this transformation by embracing digital creativity through their AI-generated campaign featuring their iconic "Pink Bambino Bag" atop a bus driving through the streets of Paris, blurring lines between art, fashion, and digital surrealism. In addition, events like Metaverse fashion week are a low-impact digital experience that showcase how brands are engaging consumers through virtual fashion, allowing global audiences to access digital fashion via NFTS and immersive platforms, thereby minimising the environmental footprint of traditional events. These developments underscore the dual transformation shaping the future of fashion.

Technology and sustainability are no longer distinct pursuits; they are intricately interwoven, each reinforcing the other. As environmental concerns escalate, the imperative for responsible operations grows stronger, and technological advancements provide the tools for fashion brands to embrace sustainable practices meaningfully. This synergy is not merely reshaping the fashion industry but charting its future, setting a new standard for innovation, accountability, and progress.

Works Cited

Birch, K. (2021). How digital assets are changing the business of fashion. [online] Businesschief.com. Available at: https://businesschief.com/technology-and-ai/how-digital-assets-are-changing-business-fashion?

Capella University (2024). Shein's Ethical Dilemma: Forced Labor and Working Conditions. [online] Bartleby.com. Available at: https://www.bartleby.com/docs/business/778812?

Dolan, L. (2022). I went to fashion week in the metaverse. [online] CNN. Available at: https://edition.cnn.com/style/article/fashion-week-metaverse-review/index.html.

Eileen Fisher Renew (2025). Eileen Fisher Renew. [online] Eileen Fisher Renew. Available at: https://www.eileenfisherrenew.com/

Gibson, K. (2024). 5 Key Benefits of Integrating AI into Your Business. [online] Business Insights Blog. Available at: https://online.hbs.edu/blog/post/benefits-of-ai-in-business.

Guinebault, M. (2022). Metaverse Fashion Week: Fashion meets Decentraland's extravagant denizens. [online] Fashion Network. Available at: https://in.fashionnetwork.com/news/Metaverse-fashion-week-fashion-meets-decentraland-s-extravagant-denizens,1392070.html.

Harper's Bazaar (2017). The History of Haute Couture. [online] Harper's BAZAAR. Available at: https://www.harpersbazaar.com/uk/fashion/fashion-news/news/a31123/the-history-of-haute-couture/.

Hughes, T. (2023). Fashion Marketing: The Jacquemus Way - Fly High Media™. [online] Fly High Media. Available at: https://www.flyhighmedia.co.uk/blog/marketing-in-luxury-fashion-the-jacquemus-way/

Igini, M. (2023). 10 concerning Fast Fashion Waste Statistics. [online] Earth.org. Available at: https://earth.org/statistics-about-fast-fashion-waste/.

Klein, J. (2022). After Metaverse Fashion Week, What Are the Implications of High Fashion Going Digital? [online] Teen Vogue. Available at: https://www.teenvogue.com/story/after-metaverse-fashion-week-what-are-the-implications-of-high-fashion-going-digital

Licata, M. and Kemp, W. (2020). The future of work in fashion. [online] UN Today. Available at: https://untoday.org/the-future-of-work-in-fashion/.

MacRae, D. (2025). Exploring the environmental impact of blockchain technology The Block. [online] The Block. Available at: https://blockchaintechnology-news.com/news/exploring-the-environmental-impact-of-blockchain-technology/.

Makos, J. (2024). What Is PESTLE Analysis? A Tool for Business Analysis. [online] PESTLE Analysis. Available at: https://pestleanalysis.com/what-is-pestle-analysis/.

McKinsey & Company (2022). What Is Industry 4.0 and the Fourth Industrial Revolution? [online] McKinsey & Company. Available at: https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-are-industry-4-0-the-fourth-industrial-revolution-and-4ir.

Miller, K. (2020). The Triple Bottom Line: What It Is & Why It's Important. [online] Business Insights Blog. Available at: https://online.hbs.edu/blog/post/what-is-the-triple-bottom-line

Mistry, M. (2023). All Aboard! Jacquemus' 3D-generated Bambino bag buses take over the streets of Paris. [online] https://peepingmoon.com/. Available at: https://www.peepingmoon.com/news-fashion/news/69606/All-Aboard-Jacquemus-D-generated-Bambino-bag-buses-take-over-the-streets-of-Paris.html.

Park, Y., Ko, E. and Do, B. (2023). The perceived value of digital fashion product and purchase intention: the mediating role of the flow experience in metaverse platforms. Asia Pacific Journal of Marketing and Logistics, 35(11). doi:https://doi.org/10.1108/apjml-11-2022-0945.

Parley (2024). ADIDAS — Parley. [online] Parley. Available at: https://parley.tv/initiatives/adidasxparley.

Pearcy, A. (2021). H&M greenwashing is 'disguising the reality' of fast fashion. [online] Big Issue. Available at: https://www.bigissue.com/news/environment/hm-greenwashing-is-disguising-the-reality-of-fast-fashion/

Roy, S. and Iacoviello, A. (2024). Business Sustainability Examples | IBM. [online] Ibm.com. Available at: https://www.ibm.com/think/topics/sustainability-in-business-examples.

Saunders, J. (2025). Secrets of the 'Shein village' where KIDS were caught working in sweat shops… & why you should think twice... [online] The Scottish Sun. Available at: https://www.thescottishsun.co.uk/fabulous/14162741/shein-village-china-factories-rights/

Sky Society (2024). The Genius Jacquemus Marketing that Took Over Social Media. [online] Skysociety.co. Available at: https://www.skysociety.co/blog/The%20Genius%20Jacquemus%20Marketing%20that%20Took%20Over%20Social%20Media%20And%20Our%20Hearts

Spiliakos, A. (2018). What does 'sustainability' mean in business? [online] Harvard Business School Online. Available at: https://online.hbs.edu/blog/post/what-is-sustainability-in-business.

tacticone (2024). Case Study: Jacquemus' Innovative Marketing – A Blueprint for Modern Luxury. [online] Tacticone.co. Available at: https://www.tacticone.co/blog/jacquemus-innovative-marketing-for-modern-luxury

The Investopedia Team (2023). Macroeconomics. [online] Investopedia. Available at: https://www.investopedia.com/terms/m/macroeconomics.asp.

Thompson, C. (2023). Metaverse Fashion Week: 70 Brands Do Their Best to Showcase Style in Decentraland. [online] CoinDesk. Available at:

https://www.coindesk.com/tech/2022/03/24/metaverse-fashion-week-70-brands-do-their-best-to-showcase-style-in-decentraland

UNEP (2021). Indicator 12.6.1. [online] UNEP - UN Environment Programme. Available at: https://www.unep.org/topics/sustainable-development-goals/why-do-sustainable-development-goals-matter/goal-12.

United Nations (1987). Sustainability. United Nations. [online] Available at: https://www.un.org/en/academic-impact/sustainability.

United Nations (2025). Goal 12: Responsible Consumption and Production. [online] The Global Goals. Available at: https://www.globalgoals.org/goals/12-responsible-consumption-and-production/.

UNSW (2023). How can technological changes affect a business? | UNSW Online. [online] studyonline.unsw.edu.au. Available at: https://studyonline.unsw.edu.au/blog/how-can-technological-changes-affect-business.

Wallace, K. (2025). How do I find the right influencers to promote my company? [online] Thetimes.com. Available at: https://www.thetimes.com/business-money/entrepreneurs/article/how-do-i-find-the-right-influencers-to-promote-my-company-enterprise-network-rqd75ds83

Williment, C. (2025). Stella McCartney, Sustainability, Ethics & Fashion. [online] Sustainabilitymag.com. Available at: https://sustainabilitymag.com/articles/stella-mccartney-stays-as-ambassador-post-lvmh-share-buyback.

Zangana, P. (2022). Renovi Launches Virtual Stores for DKNY, Diageo and GM Motors During Metaverse Fashion Week. [online] Business Wire. Available at: https://www.businesswire.com/news/home/20220324005746/en/Renovi-Launches-Virtual-Stores-for-DKNY-Diageo-and-GM-Motors-During-Metaverse-Fashion-Week.

Hamiltonian Systems And The Poincaré Recurrence Theorem By Adarsh Chauhan

Abstract

The Poincaré Recurrence Theorem is one of the most non-intuitive concepts in classical physics and mathematics. This paper provides a fresh perspective on the theorem and Hamiltonian systems. First proposed during Poincaré's work on the Three-Body Problem, the recurrence theorem states that any finite, isolated, energy-conserving system will eventually return arbitrarily close to its starting state, and keep doing so forever. While the theorem checks out mathematically, it seems to go directly against the second law of thermodynamics, where entropy always increases. This paper looks into the reasoning behind the theorem, ties it in with Liouville's Theorem, and examines why such recurrence breaks down when applied to something as vast and complex as the universe. In the end, while the theorem may be solid on paper, its real-world relevance is limited.

Introduction

The Poincaré Recurrence Theorem states that any finite-volume, energy-preserving, isolated system, when evolved over an infinite period, will return arbitrarily close to its original state—and do so infinitely often. While this has been verified mathematically, it appears to contradict the second law of thermodynamics, which suggests entropy always increases

Poincaré first introduced this theorem while working on the Three-Body Problem, during a competition held by King Oscar II. Even though it's a solid piece of mathematics, it brings up some big questions about time, entropy, and whether the universe is truly deterministic. This paper revisits the theorem in the setting of Hamiltonian systems and explores whether recurrence really holds up when we think about the universe as a whole. While the idea works out cleanly in ideal, theoretical setups, it starts to fall apart when extended to cosmological boundaries—showing that mathematical certainty doesn't always translate to physical reality.

Results

Hamiltonian Systems

Hamiltonian systems are a fundamental class of classical mechanical systems governed by Hamilton's equations. These systems consist of a phase space $M = \mathbb{R}^{2n}$, or more generally a symplectic manifold, and a Hamiltonian function $H: M \to \mathbb{R}$. The time evolution of such systems is described by:

$$dq_i/dt = \partial H/\partial p_i, \quad dp_i/dt = -\partial H/\partial q_i \quad \text{for } i = 1, 2, ..., n$$

These equations preserve total energy and, by extension, the structure of phase space volume, making them suitable candidates for recurrence analysis.

Proof of the Poincaré Recurrence Theorem

The general proof uses measure theory. Let $(X, \mathbb{F}, \mu)$ be a finite measure space and $T: X \to X$ a measure-preserving transformation. For any measurable set $A \subset X$ with $\mu(A) > 0$, define the set of return times:

$$R_A = \{ n \in \mathbb{N} : \mu(A \cap T^{-n}(A)) > 0 \}$$

By countable subadditivity and measure invariance, it follows that for almost every point $x \in A$, $T^n(x) \in A$ for infinitely many n.

Liouville's Theorem

Liouville's Theorem asserts that the volume of phase space occupied by an ensemble of Hamiltonian systems remains constant over time. For instance, a set of frictionless pendulums evolving in phase space will stretch and rotate as a cloud of points, but the volume of that cloud will not change. This conservation of phase space volume is essential to proving recurrence in Hamiltonian systems.

Discussion

Recurrence and Cosmology

The application of the recurrence theorem to cosmology introduces deep conceptual challenges. If the universe behaves like a finite, isolated Hamiltonian system, recurrence would suggest eventual return of states—even our current configuration. However, this conflicts with the second law of thermodynamics, which suggests a unidirectional arrow of time.

Two primary possibilities arise:

1. The universe is finite and isolated, implying eventual recurrence.
2. The universe is not a closed system, or its boundaries prevent recurrence, implying entropy increase and temporal irreversibility.

These interpretations question whether time is truly linear or a looping structure, and whether entropy is a statistical artifact rather than a universal law. Even though Liouville's Theorem holds locally, its global application to the universe is speculative and theory-dependent.

Implications and Limitations

While the Poincaré Recurrence Theorem provides elegant insights into reversible mechanics, its limitations must be acknowledged. Real-world systems encounter decoherence, entropy production, and chaos, which reduce the likelihood of recurrence on observable

timescales. Moreover, cosmological systems are rarely truly isolated, and include phenomena such as dark energy, black holes, and expansion that lie outside the theorem's scope.

Materials And Methods

This work is theoretical in nature and does not involve experimental methods. The recurrence theorem and Liouville's theorem were studied using foundational texts in classical mechanics and measure theory. The phase space dynamics were modeled using standard analytical techniques, and recurrence conditions were derived mathematically.

Proofs and interpretations were constructed based on:

- Hamilton's canonical equations
- Standard measure-preserving dynamics
- Classical thermodynamic laws and their interpretations

Acknowledgements

I would like to thank my teachers and my school for enabling and supporting me throughout my research journey.

Works Cited

Arnold, V. I. (1989). Mathematical Methods of Classical Mechanics (2nd ed.). Springer.

Oxtoby, J. C. (1971). The Poincaré Recurrence Theorem. In Measure and Category (pp. 124–130). Springer.

Penrose, R. (2016). Fashion, Faith, and Fantasy in the New Physics of the Universe. Princeton University Press.

Goldstein, H., Poole, C. P., & Safko, J. L. (2002). Classical Mechanics (3rd ed.). Addison-Wesley.

Earman, J. (2006). The "Past Hypothesis": Not even false. Studies in History and Philosophy of Modern Physics, 37(3), 399–430.

Barrow-Green, J. (1997). Poincaré and the Three Body Problem. American Mathematical Society.

AI & Personalized Immunology: Predicting Immune Reactions to Vaccines - A Systematic Review By Shobhita Nehra

Abstract

Background

Artificial Intelligence (AI) is increasingly being integrated into immunology to enhance the understanding of complex immune responses. By leveraging machine learning and predictive modeling. AI opens new pathways to analyze large-scale immunological data that are not available through traditional methods. A key application of AI is personalized medicine. AI can be integrated in this field for the prediction of individualized immune reactions, contributing to immunization strategies unique to each patient's biology.

Objective

The aim of this systematic review is to explore how AI can predict patient-specific immune reactions to vaccines for vaccine responses and efficiency in emerging diseases, minimize adverse effects, and transform vaccine development.

Method

Two rounds of screening were conducted on PubMed. The first phase involved an examination of only the title and abstract, followed by a full-text assessment of the paper identified by my search strategy. The included studies involved homo sapiens and examine AI applications in immunology, specifically personalized medicine, with no geographical and disease restriction. Then the PRISMA flow chart to document the review process, remove duplicate papers, and track study identification and final outcomes.

Results

Among 371 papers retrieved from my search strategy, 347 were screened for title and abstract, from which only 51 were sought for retrieval of full-text screening. In the end, only 31 papers met the eligibility criteria of the review. This research demonstrates the effectiveness of advanced AI models in predicting personalized immune responses to vaccines. After assessing a range of models, Routed LSTM and Vaxingn-ML consistently outperformed traditional approaches and other machine learning systems. Different biomarkers were analyzed to associate with immune responses and their effect on model precision. Personalized variables such as age, gender, comorbidities, and prior immune health all influence model outputs. AI-driven predictions are more precise and stronger than clinical trial data.

Conclusion

The findings in this study suggest that AI not only allows for more precise and accurate responses but also provides insightful understandings of immunological factors. It helps pave the

way for tailored vaccine designs and strategies by exploring the applications in personalized immunology.

Introduction

One Vaccine. Millions of bodies. Infinite immune outcomes. Vaccines have revolutionized modern medicine, averting an estimated 4 million deaths annually.[1] While representing one of medicine's greatest triumphs, their success has been constrained by one critical flaw: they are designed for populations, not individuals. Vaccines have traditionally followed a one-size-fits-all approach despite a world where medical breakthroughs unfold rapidly. COVID-19 exposed the limitations of this process. While mRNA vaccines achieved 95% efficacy in trials, real-world efficacy plummeted to 35-45% in immunocompromised patients.[2]

No two immune systems are alike. Factors such as age, genetics, environmental exposure, microbiome composition, and immune history create a unique immunological system for every individual. This complexity has been a barrier to personalization, but also the very reason we need it. We live in an era demanding precise medicine. By identifying unique patterns of cell behavior and gene expression, scientists aim to tailor interventions for maximum efficacy and minimize adverse effects. This is where AI offers a transformative solution. AI is rapidly becoming an integral force in healthcare. AI, specifically machine learning, has proven to be capable of managing and interpreting vast biological datasets. It can synthesize information from genomics, transcriptomics, proteomics, and clinical records. These technologies have the potential to model immune behavior, response to vaccines, and flag individuals at risk of adverse effects.

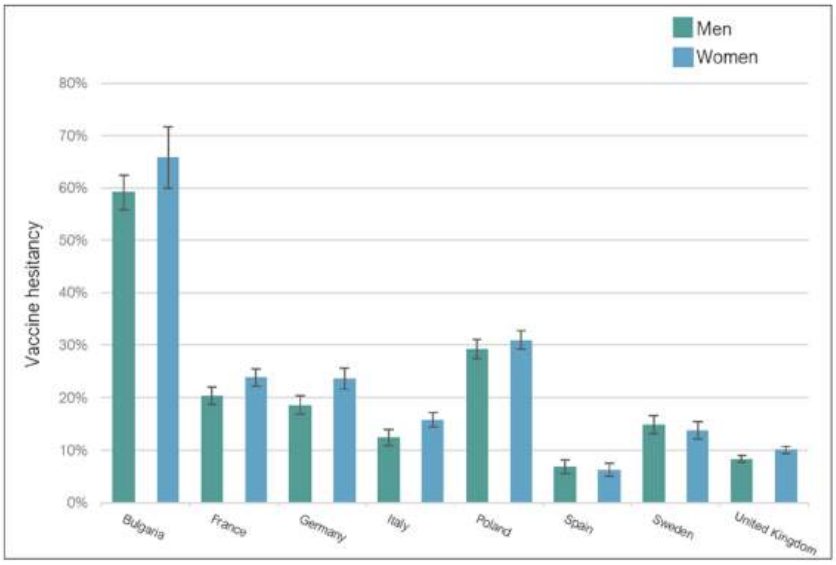


Figure 1. Vaccine Hesitancy (COVID-19) in 8 Different European Countries
Source: Steinert, Janina I., et al. Science Advances, Apr. 2022

[1] "Vaccination and Immunization Statistics." *UNICEF DATA*, 9 July 2019, https://data.unicef.org/topic/child-health/immunization/.
[2] Link-Gelles, Ruth. "Interim Estimates of 2024–2025 COVID-19 ..." *MMWR. Morbidity and Mortality Weekly Report* 74 (February 27, 2025). https://doi.org/10.15585/mmwr.mm7406a1. .

Despite these promises, AI's applications remain largely understudied. As pathogens evolve faster than ever, global vaccine hesitancy increases. Figures 1 and 2 highlight how, despite the severity and urgency of COVID-19, resistance to vaccination was prevalent in many countries around the world. We can no longer afford a trial-and-error approach. While AI is successful in domains such as diagnostic imaging and drug discovery, it remains underutilized in vaccine responses. This lack of personalized prediction can lead to suboptimal efficacy, unanticipated side effects, and reduced public confidence in immunization.

AI's real power lies in its adaptability. In an ever-evolving world, these systems learn continuously and can be updated as new data emerges in real time. Deep learning and probabilistic models can be trained to detect patterns and predict outcomes at the individual level. This would help in identifying how a person's immune system is likely to react to a specific vaccine based on their unique biological profile. These insights can transform vaccine formulation, dosing, and applications. Some people, such as those with HIV or organ transplant patients, don't respond well to vaccines. Personalized vaccines are required by these high-risk groups as current models don't capture the full immune response. Further, it would help in the development of cancer vaccines as they need to target the right tumor parts and deal with diverse patient responses. It would also help eliminate issues such as long production times of vaccines, complex processes, and testing cycles. It would also help in fully understanding the immune system, which is also a limiting factor in improving vaccines.

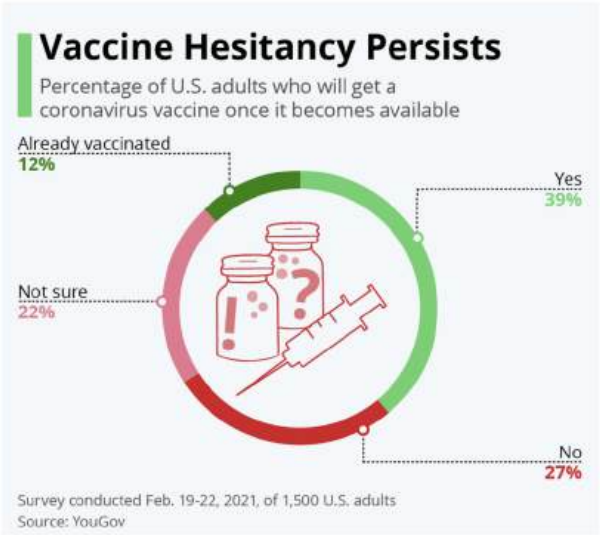


Figure 2. Vaccine Hesitancy (COVID 19) in United States
Source: Roper, Willem. *Statista*, 23 Feb. 2021,

This systematic review aims to unpack the current and emerging uses of artificial intelligence in predicting patient-specific immune responses to vaccines. It examines the AI model training and techniques, biomarker analysis, the factors affecting personalized immune response prediction, vaccine responses, comparison of AI with clinical data, and case studies.

Methodology

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines. A comprehensive search was conducted using the PubMed database to identify relevant literature on artificial intelligence and vaccines. PubMed was selected due to its extensive biomedical literature and reliable studies. For this comprehensive search, I created a search strategy as follows: ("vaccin*"[Title] AND ("Artificial Intelligence"[MeSH Terms] OR "Artificial Intelligence"[Title/Abstract])) AND (y_5[Filter]). The search was structured, combining controlled vocabulary (MeSH terms) and text keywords to maximize precision and retrieve all applicable papers. This strategy focuses on research published within the last 5 years to capture recent and significant developments in the field, as it is an upcoming branch in medicine. The keywords were developed around the core concepts of artificial intelligence and vaccination to ensure focus and relevance. From this search, a total of 371 papers were retrieved initially.

The eligibility criteria of the studies focused on the population group of *Homo sapiens*. The research had to address the intersection and core themes of immunology, artificial intelligence, and personalized medicine. There were no geographical, study design, article type, or language restrictions. There was a 5-year restriction on the studies retrieved to ensure the research was recent and suitable. Studies were mainly evaluated based on their context and research. Papers that had non-patient-specific data were excluded due to their irrelevance.

Once the applicable papers were retrieved, a 2-stage screening protocol was followed to ensure a rigorous and systematic selection of studies, enhancing the reliability and precision of the final literature. The first round of screening included only reading the title and abstract, followed by a full-text screening. This approach allowed for efficient management of the large records of data retrieved in the initial search, it helped to balance breadth and depth, as it allowed for an initial broad review followed by a more comprehensive review.

After the initial search on PubMed, which yielded 371 studies, duplicates, and irrelevant studies were removed, resulting in 347 studies eligible for title and abstract screening. This stage was a preliminary filter to assess the relevance and fit of each study based on its focus and implications. Studies that were outside of the scope and not of relevance were excluded.

From the 347 studies screened for title and abstract, 51 were selected for full-text screening, allowing for a more detailed evaluation. This helped me confirm that the studies chosen were of use and had sufficient data to include in the final data set. 39 full texts were successfully retrieved and reviewed, and the remaining 12 papers were unavailable for full-text screening. From this set, 6 studies were excluded as they did not include patient-specific applications, and 3 studies were removed as they were not related to predictive AI.

Ultimately, 31 studies met the inclusion criteria and were retrieved for final synthesis. This entire selection process was documented using a PRISMA flow diagram. The PRISMA flow diagram serves to depict numerical results derived from the reviews as well as the inclusion

decision process. My flow chart effectively demonstrates the steps involved in obtaining the appropriate research and final analysis. It helps to facilitate and streamline the extraction of data.

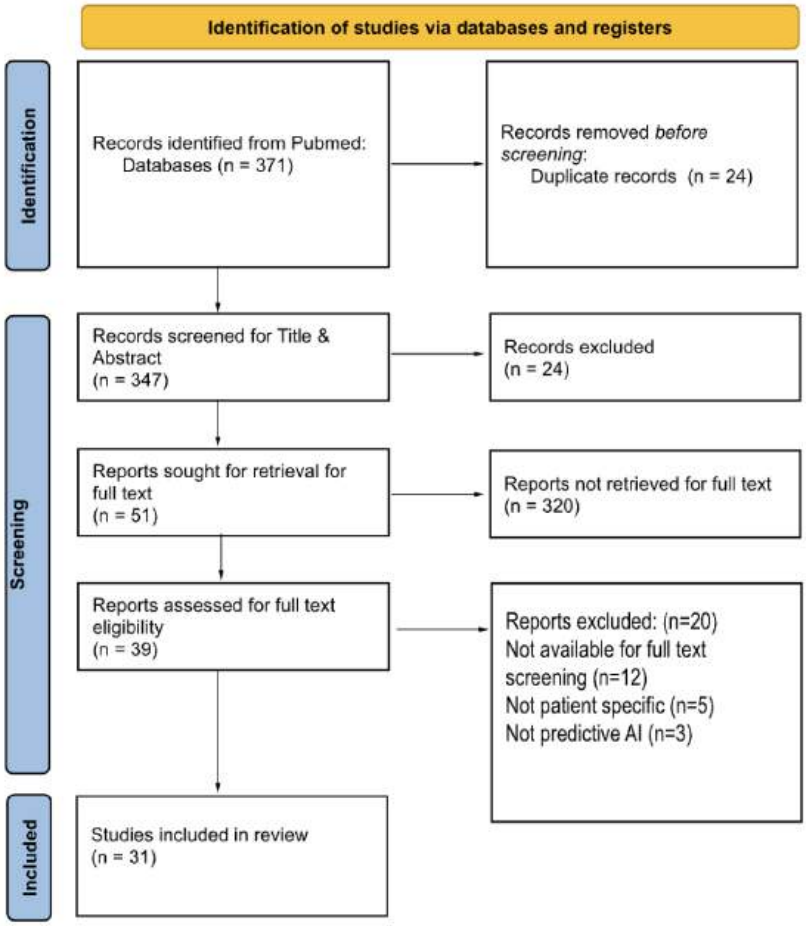


Figure 3. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the study selection process

AI Model Training and Techniques

Artificial Intelligence (AI) refers to the simulation of human intelligence in machines. AI models in immunology can help predict individual immune responses to vaccines by integrating various variables such as biomarkers, patient-specific variables, and more. Through this, AI bridges immunology and innovation.

From the final dataset, a wide array of artificial intelligence and machine learning models were exploring longitudinal changes in immune function, such as the development of immunosuppression or the activation of certain immune cells and evaluated to predict immune responses to vaccines, with a focus on implications in personalized immunology.

To begin with, logistic regression (LR), one of the earliest tools used in predictive modeling, has proven useful in binary classification tasks such as predicting whether a vaccine will induce a sufficient immune response. Its strength lies in its interpretability. Clinicians can

trace how input variables like age, immune history, and comorbidities impact outcomes. However, LR assumes linear relationships and can underperform with complex, non-linear immunological data. (Azarfar et al. 2025) (Jamshidi et al. 2023) (Ong et al. 2020)

Next, random forests improve upon this by constructing multiple decision trees on subsets of the data and aggregating their predictions. This approach captures non-linear interactions among features, such as the interplay between cytokine levels and antibody titers. This model can highlight the most influential factor behind immune variability. However, while these models are accurate, they may lack transparency compared to simpler methods, posing challenges in clinical application. (Montesi et al. 2024)

Long Short-Term Memory (LSTM) networks are a subclass of recurrent neural networks (RNNs) designed to handle sequential data. LSTMs can analyze longitudinal immune data such as changes in B-cell and T-cell populations over time. They can be used to represent dynamic immune responses after multiple doses of vaccinations because of their internal memory processes, which enable them to maintain critical patterns over time steps. Routed LSTM expands on LSTM by introducing an adaptable mechanism that routes each input via various LSTM "experts" according to its content. For personalized immunology, where patient data differs greatly, this makes it very potent. For example, older patients' immune profiles might exhibit distinct patterns compared to younger patients', routed LSTMs dynamically allocate them to distinct sub-networks, simulating personalized immune processing. (Azarfar et al. 2025)

Furthermore, deep generative models known as variational autoencoders (VAEs) are capable of producing new samples by learning compressed, continuous representations of input data, or latent spaces. VAEs have been used in immunology to mimic immune responses to novel vaccine formulations and model T-cell and B-cell receptor sequences. VAEs can forecast the potential responses to antigens of individuals with certain Human Leukocyte Antigen (HLA) alleles or immunological backgrounds by learning latent features that reflect immune variability. This is especially important when developing cancer vaccines because tumor neoantigens differ greatly from patient to patient. VAEs facilitate the development of patient-specific vaccination regimens by enabling the simulation of appropriate immune responses and the construction of hypothetical neoantigen profiles. They are an effective tool in the prediction of personalized immunology because of their capacity to extrapolate to unknown biological circumstances. (Matson et al. 2024)

Moreover, reverse vaccinology and machine learning are combined in Vaxign-ML, which analyzes proteome and genomic data to forecast potential vaccine candidates. It assesses possible targets using characteristics like subcellular localization, adhesin frequency, and antigenicity. Its machine-learning backbone allows adaptation to patient-specific antigen recognition, particularly in infections with varying expression (Ong et al. 2020) **(**Tu, Zhiwei, et al. 2024) **(**Lv, Hao, et al. 2021) patterns among hosts, even though its main usage is in pathogen-focused vaccine creation. DeepVacPred integrates sequence embedding, reverse vaccinology, and deep learning. It predicts antigenic areas that are likely to elicit immunological responses using convolutional and recurrent neural networks. DeepVacPred can be improved with individual major

histocompatibility complex (MHC) profiles or immunopeptidomics data, even though it was mostly trained on public datasets. This helps to close the gap between generic predictions and customized vaccination approaches. (Yang, Zikun, et al. 2021)

In addition, a multi-omics integration platform called DIABLO (Data Integration Analysis for Biomarker Discovery utilizing Latent Components) was created to link features across the immunological, proteomic, transcriptomic, and genomic levels. DIABLO makes it easier to find cross-omic indicators that affect immunogenicity in the context of vaccine response prediction, such as genes associated with antigen processing or post-transcriptional changes that impact epitope presentation. Such integration is beneficial to personalized immunology because it clarifies why transcriptome or proteomic changes might cause individuals with comparable genotypes to exhibit different vaccination responses. For example, DIABLO has been used to stratify patients before vaccination by predicting the relationship between cytokine gene expression and unfavorable vaccine reactions. (Aevermann, Brian D., et al. 2021)

By using amino acid sequences to predict protein 3D conformations with almost experimental accuracy, AlphaFold has transformed structural biology. AlphaFold aids in the modeling of antigen architectures, evaluation of epitope accessibility, and prediction of antibody-antigen binding affinity in vaccination research. When creating customized vaccines, these structural insights are especially important because cancer immunotherapy requires optimizing antibody interactions with distinct tumor antigens. AlphaFold allows for the in silico assessment of immunological compatibility by simulating the antigen and anticipated antibody response from a person's immune repertoire. This enables scientists to choose epitopes that are expected to be stable, exposed, and highly immunogenic for a particular person, personalizing vaccines at the structural level. (Ong et al. 2020)

Lastly, Genentech's PIONEER TM uses artificial intelligence (AI) to analyze and categorize immune responses in a variety of illness situations. It combines immunological, genetic, and clinical data to identify non-responders or customize vaccine recommendations. Its emphasis on immunological modeling across populations makes it a contender for scaling tailored predictions based on a variety of real-world data, even though specifics are secret. (Mørk, Sofie Kirial, et al. 2022)

To conclude, the technological foundation of a new era in immunology, one that emphasizes personalized responses over general population data, is formed by these models taken together. The methods examined provide complementary strengths, ranging from interpretable logistic regression models to complex VAEs and LSTMs. Some of these tools prioritize explainability and feature importance, while others are particularly good at identifying high-dimensional patterns and making predictions in the face of ambiguity. Crucially, their combined use in vaccine response modeling is expanding the realm of personalized immunology's potential, allowing for vaccines that are not only safe and efficient but also tailored to the immune system of each individual.

AI Model	Model Type	Primary Use	Strengths	Limitations	Personalized Capability	Example Application
Logistic Regression (LR)	Traditional ML (Linear)	Binary classification of immune responses	High interpretability; simple	Assumes linearity; poor with complex data	Low	Predicting vaccine efficacy using age/comorbidity
Random Forests	Ensemble ML (Decision Trees)	Identifying key immune predictors	Captures non-linear interactions; feature importance	Less transparent; “black box”	Moderate	Evaluating the influence of cytokine/antibody patterns
LSTM	Deep Learning (RNN)	Time-series modeling of immune response	Handles sequential data; learns temporal patterns	Needs large datasets; hard to interpret	High	Modeling changes in immune cell levels post-vaccine
Routed LSTM	Advanced Deep Learning (RNN variant)	Personalized routing of immune profiles	Dynamic sub-networks; better personalization	Complex to train; low transparency	Very High	Tailoring vaccine response prediction for age groups
VAE (Variational Autoencoder)	Deep generative model	Latent immune response representation	Learns hidden patterns; extrapolates to novel inputs	Complex; interpretability limited	Very High	Forecasting responses to cancer vaccines
Vaxign-ML	Reverse Vaccinology + ML	Predicting vaccine targets from pathogen data	Integrates proteomic/genomic features; pathogen focus	Less personalized; trained for general cases	Modeate	Selecting antigens based on adhesion/localization
DeepVacPred	Deep learning + sequence embedding	Predicting antigenic regions	Combines CNN, RNN, and reverse vaccinology	Requires public data; personalization depends on inputs	High (with personal data)	Predicting epitopes using MHC/immunopeptidomics
DIABLO	Multi-omics integration ML	Biomarker discovery for vaccine response	Cross-omics insights identify hidden drivers	Requires multiple datasets; complex interpretation	Very High	Stratifying patients based on cytokine transcriptomics
AlphaFold	Deep learning (structural bio)	Protein structure prediction	High structural accuracy; aids epitope analysis	Focused on structure only, not temporal	High	Modeling tumor antigen-antibody interactions
PIONEER (Genentech)	Proprietary AI platform	Real-world immune response classification	Multi-modal data integration, scalable	Not fully disclosed; transparency limited	Very High	Identifying vaccine non-responders in clinical settings

Figure 4. Summary of AI Models

Biomarker analysis

Biomarkers are measurable biological indicators that provide information about processes in the body. They can be molecules, genes, immune cells, proteins, etc, which reflect a person's health status or response. In immunology, biomarkers can help assess how effectively the immune system responds to a vaccine.

The increasing use of AI in biomarker identification is proving to be very beneficial, especially for vulnerable populations like people with HIV or chronic illnesses (Montesi et al. 2024) (Cotugno, Nicola, et al. 2020). Refined forecasts of infection status and vaccination effectiveness are made possible by predictive scores of immunogenicity, which are constructed using gene expression and serological data. Future studies must confirm and build upon these findings across a range of groups because the incorporation of biomarkers into personalized medicine presents a viable avenue for vaccine development and application.

Anti-receptor-binding domain (anti-RBD) antibodies are important biomarkers for forecasting solid organ transplant recipients' vaccination responses. Based on a patient's anti-RBD antibody levels, they were classified as responders or nonresponders; a response was defined as ≥0.8 U/mL and a lack of a response as <0.8 U/mL. This cutoff point functions as a useful gauge for assessing immunological effectiveness following immunization. Crucially, studies discovered that, in contrast to low-responders and nonresponders, high-responder patients showed a wider variety of factors affecting their immune responses, suggesting increased complexity and diversity in immunogenicity. (Azarfar, Ghazal, et al. 2025)

Beyond this, numerous more biomarkers connected to vaccine efficacy have been found by other researchers. Non-neutralizing antibodies (non-NAbs) that target the gp120-V1V2 region were found to correlate with protection, indicating that CREB1 expression and its downstream gene targets are linked to antibody responses to V1V2 epitopes. The results of macaque investigations that showed immunological signatures suggestive of vaccine efficacy further corroborated the conclusion that these V1V2-specific antibodies were the best predictors of protection. However, because these correlates vary throughout research and populations, their validation is still difficult. (Klasse, P. J. 2022)

Machine learning techniques have been applied to predict humoral responses in HIV-infected patients based on a variety of indicators, such as Body Mass Index (BMI), CD4 T-cell count, CD4/CD8 ratio, and previous SARS-CoV-2 infection status (Montesi et al. 2024). These indicators had a major impact on antibody production, indicating their value in tailoring immunization campaigns for immunocompromised populations. Likewise, in more general biomedical settings, biomarkers including Granzyme B (GrB), T-bet, and FoxP3 have been used to measure Th1/Treg cell responses and cytotoxic lymphocyte activity, particularly in research on the effectiveness of cancer vaccines (Mirsanei, et al. 2020). These biomarkers at the transcriptional and protein levels shed light on how the immune system works and how well treatments work.

The baseline dendritic cell signatures for Hepatitis B Virus (HBV) vaccine response prediction, where a ratio between NDRG1-expressing mDC2 and CDKN1C-expressing mDC4

cells functioned as a predictive tool, are among the other noteworthy biomarkers found in the research (Aevermann, Brian D., et al. 2021). Immune profiles that could be trustworthy pre-vaccination markers were also discovered using multi-omics techniques. High-resolution proteome microarray research and machine learning models such as multitask support vector machine (SVMs), have shown that Pf-specific antigens are biomarkers for predicting sterile protective immunity in the context of malaria immunization. (Wistuba-Hamprecht, Jacqueline, et al. 2024)

Research on cancer biomarkers has grown dramatically, and they are now categorized as genetic, protein, imaging, and liquid biomarkers, all of which provide targeted, non-invasive insights into the biology of tumors. Prostate-specific antigen, or PSA, is used as a marker in the monitoring of prostate cancer treatment, and artificial intelligence techniques improve the accuracy of biomarker predictions for tailored treatments (Kumar, Anant, et al. 2024). Furthermore, research has employed AI to find biomarkers for SARS-CoV-2 infection and lupus nephritis, such as RPL10 gene expression and the presence of the HLA-B15 allele, which both affect antibody levels and vaccine responses (Li, Jing, et al. 2023). In addition to reports of increased T-cell counts in children who received vaccinations, a nine-transcript signature was put forth as a predictor of vaccine protection. These results highlight the usefulness of transcriptome analysis in identifying immunological changes brought on by vaccines.

Biomarker	**Type**	**Measurement**	**Associated Condition**	**Use in Vaccine Prediction**	**Limitation**
Anti-RBD Antibodies	Serological	Serum	Solid organ transplant recipients	Classifies responders vs non-responders post-COVID-19 vaccination (≥0.8 U/mL)	Cutoff may not generalize across populations
Non-neutralizing antibodies (V1V2-specific)	Serological	Plasma	HIV patients, macaques	Correlates with protection; associated with CREB1 gene expression	Correlation varies by population and study
CREB1 & downstream genes	Transcriptomic	Gene expression profiles	HIV vaccine response studies	Linked to antibody responses to V1V2 epitopes	Requires transcriptome data; population-specific
BMI, CD4 count, CD4/CD8 ratio, prior infection	Clinical & Immunological	Patient records & blood tests	HIV-infected patients	Predicts humoral vaccine responses	May vary with region and comorbidity profiles
Granzyme B, T-bet, FoxP3	Protein / Transcriptomic	PBMCs	Cancer immunotherapy research	Measure cytotoxicity, Th1/Treg cell balance	Requires flow cytometry or transcriptomics

NDRG1/CDKN1C mDC2/mDC4 ratio	Cellular (Dendritic Cells)	Immune profiling	HBV vaccine recipients	Pre-vaccination biomarker for immune response	Needs specialized cell-type markers
Pf-specific antigens (Malaria)	Proteomic	High-resolution proteome microarray	Malaria-endemic populations	Predict sterile protective immunity	Context-specific to malaria vaccines
Prostate-Specific Antigen (PSA)	Protein	Blood serum	Prostate cancer patients	Monitors treatment efficacy; AI improves prediction	Specific to prostate cancer, not vaccine-related
RPL10 expression, HLA-B15 allele	Transcriptomic & Genetic	PBMCs, Genotyping	SARS-CoV-2, Lupus nephritis	Affect vaccine antibody response levels	May require genomic sequencing
T-cell counts (post-vaccination)	Cellular	Blood	Pediatric vaccine recipients	Correlates with stronger immunogenicity	Basic marker; not mechanistic alone
Nine-transcript signature	Transcriptomic	Whole blood RNA	General vaccine prediction	Predicts vaccine protection	Needs validation across vaccines/populations

Figure 5. Biomarkers Associated with Vaccine Responses

Vaccine Responses

The dataset's findings add to the increasing amount of research showing how useful artificial intelligence (AI) is for forecasting vaccine-induced immune responses and directing precision immunology. The routed LSTM model outperformed conventional models in predicting SARS-CoV-2 vaccination responses in solid organ transplant recipients (SOTRs), with an MSE of 0.02 and a Pearson r of 0.79. (Azarfar et al. 2025)

In a variety of models, other machine learning algorithms also exhibit promise. For example, a model that predicts humoral responses in people living with HIV (PLWH) identified infection history, BMI, and CD4+ T-cell count as important factors, suggesting that booster dosages may be able to compensate for immunogenicity deficiencies (Montesi et al. 2024). The significance of baseline immunological signatures in response variability was further supported by studies that successfully predicted hepatitis B vaccine responses based on dendritic cell profiles by combining single-cell RNA sequencing and feature selection approaches. (Aevermann, Brian D., et al. 2021)

Mechanistically, the creation of peptide-based vaccines against viruses such as SARS-CoV-2 and poxviruses has been sped up using AI-enabled reverse vaccinology and epitope prediction systems like PoxiPred (Martinez, Gustavo Sganzerla, et al. 2024). AI's involvement in precision oncology has been confirmed by the safety and long-lasting T-cell responses shown by personalized neoantigen vaccines (Mørk, Sofie Kirial, et al. 2022).

Additionally, deep learning has become a game-changing tool for understanding viral evolution, predicting population vulnerability, and optimizing antigen selection, all of which are especially pertinent to influenza and malaria.

Factors affecting personalized medicine response prediction

The effectiveness of personalized vaccines relies on interactions among biological, clinical, and demographic factors. Research consistently indicates that age is a significant determinant. Younger age groups often face a higher incidence of adverse effects, while older age groups, specifically those who have undergone transplants or comorbid conditions, frequently show reduced antibody responses (Azarfar, Ghazal, et al. 2025). Other influential factors include sex, body mass index (BMI), race, genetic markers like HLA alleles, specific gene expression patterns, previous infections (notably SARS-CoV-2), immune health metrics such as CD4/CD8 ratios, and the use of immunosuppressive medications (particularly mycophenolate mofetil and tacrolimus) **(**Meng, Lingxin, et al. 2024) **(**Singh, Prateek, et al. 2022). Specifically, lower dosages of mycophenolate and decreased levels of tacrolimus were associated with enhanced immune responses. In immunocompromised groups, such as individuals with HIV and cancer patients, factors such as the duration during transplantation, T-cell subsets, and the presence of immune memory play a crucial role in determining outcomes. (Azarfar, Ghazal, et al. 2025)

To address the intricacies of biological systems, researchers are utilizing artificial intelligence and machine learning techniques capable of identifying non-linear relationships and interactions with extensive, multi-parameter datasets. Approaches such as random forests, ADA boost (adaptive boosting), ensemble models, and artificial neural networks (ANNs) have been employed to forecast immune responses by synthesizing features including omics data, immune biomarkers, and clinical variables. These models assist in categorizing patients as responders or non-responders, identifying patterns of neoantigen presentation, and enhancing peptide design for cancer vaccines. Research indicates that the inclusion of data from dendritic cell subsets, transcriptional profiles, and mutation-specific T-cell responses markedly enhances prediction accuracy. However, in certain instances, the predictive significance of specific factors, like antibody levels or gene signatures, may decline with repeated vaccine doses. (Kumar, Anant, et al. 2024)

Although advancements have been made, numerous obstacles persist in achieving completely accurate and personalized response predictions **(**Hederman, Andrew P., and Margaret E. Ackerman 2023). The presence of small, noisy datasets in immunology hampers the reliability of deep learning models. Additionally, factors such as tumor heterogeneity, immune imprinting, and gene-environment interactions add complexity to these predictions **(**Kumar, Anant, et al. 2024) **(**Meng, Lingxin, et al. 2024). Furthermore, issues such as delays in the collection of clinical data, lack of standardization in ontologies, and the necessity for real-time adaptability limit the effectiveness of current tools (Huffman, Anthony, et al. 2022). Nevertheless, continuous progress in AI, such as the application of variational autoencoders, liquid biomarkers, and

high-throughput serological assays, presents opportunities for developing more precise, scalable, and individualized strategies in the fields of vaccinology and cancer immunotherapy. (Matson, Robert P., et al. 2024)

Comparison of personalized AI responses with clinical data

AI has led to a new era of personalized medicine in healthcare. There are many differences in the approach of AI as compared to traditional clinical data.

To begin with, AI models can follow longitudinal changes in immune function, such as the development of immunosuppression or the activation of certain immune cells, because they can examine dynamic factors across time and can be updated with new data. This allows for more precise individual-specific vaccine response predictions. It provides insights into how different biomarkers can impact vaccine efficacy on a personalized level **(**Azarfar, Ghazal, et al. 2025). In contrast, clinical data usually focuses on more general demographic traits such as age, sex, and comorbidities. Although such information is crucial for comprehending broad population patterns, it is insufficiently detailed to customize vaccination tactics to each person's unique immune profile. Clinical trials may miss crucial nuances of the immune system, such as changes in T cell subsets over time or modifications in immune response mechanisms. (Liu, Ge, Brandon Carter, and David K Gifford. 2021)

Further, the ability of AI models to incorporate multi-omics data and model intricate relationships between genetic and environmental factors is the foundation of their potential. These AI models provide a degree of personalization that conventional clinical models are unable to match by identifying important biomarkers or immune cell signatures that function as predictors of vaccine success **(**Aevermann, Brian D., et al. 2021). Clinical studies, on the other hand, are constrained by their dependence on large cohorts and generalized responses. Individual differences are frequently not taken into account by these trials, despite the fact that they are crucial for assessing vaccine safety and efficacy across a community. **(**Hederman, Andrew P., and Margaret E. Ackerman. 2023) **(**Cotugno, Nicola, et al. 2020)

Moreover, the capacity of AI models to use data from small, specialized cohorts to forecast vaccine responses for particular populations is one of their main advantages. These models can spot trends that are frequently overlooked in more comprehensive clinical data, like the relationship between vaccine efficacy and particular dendritic cell states or tumor neoantigens. AI can also combine information from machine learning-based simulations or in vitro testing, adding layers of understanding about vaccine reactions that conventional clinical data cannot **(**Cotugno, Nicola, et al. 2020).

AI methods do, however, have certain drawbacks. For training, they need access to sizable, high-quality datasets, and the quantity and quality of data they use frequently affect how predictive they are. The predictions might not be correct if the datasets are skewed or lacking. Furthermore, AI models could have trouble with rare immune responses or limited sample numbers, especially when used on certain subgroups like individuals with rare comorbidities or genetic abnormalities. Clinical data, on the other hand, can offer a wider range of information,

including sizable and varied populations that might not be adequately represented in the datasets used to train AI models. Numerous real-world factors that may affect vaccine reactions can also be taken into consideration in clinical studies.

Discussion

AI is currently a vital component of the developing field of personalized immunology. Conventional vaccine development has frequently ignored diversity brought on by genetic, environmental, and immunological factors in favor of population averages. By investigating the evaluation of AI and machine learning models such as Logistic Regression, Random Forests, Routed LSTM, DeepVacPred, Vaxign-ML, and AlphaFold to forecast individual-specific immune responses to vaccines, this study sought to close that gap. From these models, routed LSTM achieved the highest predictive accuracy. Numerous biomarkers were found to be important predictors of vaccine response, including cytokine levels, HLA type, antibody titers, T-cell responses, and gene expression patterns. Furthermore, it was noted that comorbidities, age, sex, genetic variability, and immune history were important determinants of an individual's immune response to vaccination. AI-driven models could predict vaccine reactions with more detail where traditional methods frequently fail, according to comparisons with traditional clinical data.

An interesting discovery was the higher performance of hybrid ensemble approaches, which fused deep learning and feature interpretation to provide both scientific knowledge and forecast accuracy. Influential factors such as age, comorbidities, and past infections unexpectedly displayed larger non-linear interactions with vaccine response than first thought. For instance, earlier COVID-19 exposure resulted in increased T-cell memory responses but decreased antibody responses to later vaccinations; the AI models correctly reflected this phenomenon, which was frequently overlooked in earlier empirical research. A more thorough and dynamic understanding of immunological variability was made possible by this study's multi-modal, AI-integrated methodology as opposed to previous research that relied on solitary biomarkers or linear predictors.

The dataset's richness, which combines findings and outcomes from other studies, is one of this research's main strengths. Cross-validation against actual clinical vaccine response data was made possible by the study's methodology, which increased the findings' dependability. This study provided a thorough investigation of the science underlying customized immunology by evaluating a variety of AI models in addition to in-depth biological investigations. The study provides a solid, multifaceted understanding of vaccine response prediction by combining various frameworks, such as biomarker analysis and machine learning. The findings advance precision immunology and provide a solid basis for future medical applications. The findings are also guaranteed to be extremely pertinent and forward-looking due to the thorough synthesis of recent developments. Nevertheless, there are still limitations. Some biological factors caused bias in the training datasets. Despite their remarkable accuracy, the models are nevertheless constrained by the state of immunological research. The findings' long-term applicability is

called into question by the quick development of AI technologies and immunology research. Furthermore, an experimental study in this area is required to fill in some of the gaps in the current studies and support some of the conclusions.

Future research should concentrate on adding more layers of biological data and combining longitudinal cohort studies to increase the breadth and diversity of the databases. Additionally, the gap between clinical translation and computational prediction would be closed by experimentally validating AI-predicted vaccine candidates in clinical trials. AI has the ability to completely transform customized immunology. These results point to a future in which public health initiatives and vaccine research can become more tailored, improving protection for a range of people and eventually boosting immunological resilience worldwide.

Conclusion

To conclude, this study shows how AI has the potential to transform vaccine development by precisely forecasting individual immune responses. With the help of biomarker data and multi-omics integration, advanced models such as Routed LSTM and Vaxign-ML performed better in terms of prediction and interpretability than conventional methods. These findings open the door for next-generation vaccinations that are customized to each patient's unique immune profile, notwithstanding constraints in population diversity and data variability. The use of AI in precision medicine is only getting started. As AI techniques advance, their incorporation into clinical decision-making could revolutionize preventative medicine by enabling completely customized immunization programs, in addition to optimizing vaccine design.

Works Cited

Aevermann, Brian D., et al. "Machine Learning-Based Single Cell and Integrative Analysis Reveals That Baseline mDC Predisposition Correlates With Hepatitis B Vaccine Antibody Response." *Frontiers in Immunology*, vol. 12, Oct. 2021, https://doi.org/10.3389/fimmu.2021.690470.

Amanpour, S. "The Rapid Development and Early Success of Covid 19 Vaccines Have Raised Hopes for Accelerating the Cancer Treatment Mechanism." *Archives of Razi Institute*, vol. 76, no. 1, Jan. 2021, pp. 1–6, https://doi.org/10.22092/ari.2021.353761.1612.

Azarfar, Ghazal, et al. "Using Machine Learning for Personalized Prediction of Longitudinal Coronavirus Disease 2019 Vaccine Responses in Transplant Recipients." *American Journal of Transplantation*, vol. 0, no. 0, https://doi.org/10.1016/j.ajt.2024.11.033. Accessed 26 Apr. 2025.

Cotugno, Nicola, et al. "Artificial Intelligence Applied to in Vitro Gene Expression Testing (IVIGET) to Predict Trivalent Inactivated Influenza Vaccine Immunogenicity in HIV Infected Children." *Frontiers in Immunology*, vol. 11, Oct. 2020, https://doi.org/10.3389/fimmu.2020.559590.

Gómez-Carballa, Alberto, et al. "Host Transcriptomic Response Following Administration of Rotavirus Vaccine in Infants' Mimics Wild Type Infection." *Frontiers in Immunology*, vol. 11, Jan. 2021, https://doi.org/10.3389/fimmu.2020.580219.

Hederman, Andrew P., and Margaret E. Ackerman. "Leveraging Deep Learning to Improve Vaccine Design." *Trends in Immunology*, vol. 44, no. 5, May 2023, pp. 333–44, https://doi.org/10.1016/j.it.2023.03.002.

Huffman, Anthony, et al. "COVID-19 Vaccine Design Using Reverse and Structural Vaccinology, Ontology-Based Literature Mining and Machine Learning." *Briefings in Bioinformatics*, vol. 23, no. 4, June 2022, https://doi.org/10.1093/bib/bbac190.

Jamshidi, Elham, et al. "Personalized Predictions of Adverse Side Effects of the COVID-19 Vaccines." *Heliyon*, vol. 9, no. 1, Jan. 2023, https://doi.org/10.1016/j.heliyon.2022.e12753.

Khan, Atikur R., et al. "Distributed Lag Inspired Machine Learning for Predicting Vaccine-Induced Changes in COVID-19 Hospitalization and Intensive Care Unit Admission." *Scientific Reports*, vol. 12, no. 1, Nov. 2022, pp. 1–11, https://doi.org/10.1038/s41598-022-21969-9. Jamshidi, Elham, et al. "Personalized Predictions of Adverse Side Effects of the COVID-19 Vaccines." *Heliyon*, vol. 9, no. 1, Jan. 2023, https://doi.org/10.1016/j.heliyon.2022.e12753.

Klasse, P. J. "Reappraising the Value of HIV-1 Vaccine Correlates of Protection Analyses." *Journal of Virology*, Apr. 2022.

Kumar, Anant, et al. "Personalized Cancer Vaccine Design Using AI-Powered Technologies." *Frontiers in Immunology*, vol. 15, Nov. 2024, https://doi.org/10.3389/fimmu.2024.1357217.

Leonardelli, Lorena, et al. "Literature Mining and Mechanistic Graphical Modelling to Improve mRNA Vaccine Platforms." *Frontiers in Immunology*, vol. 12, Sept. 2021, https://doi.org/10.3389/fimmu.2021.738388.

Li, Jing, et al. “Identification of Genes Related to Immune Enhancement Caused by Heterologous ChAdOx1–BNT162b2 Vaccines in Lymphocytes at Single-Cell Resolution with Machine Learning Methods.” *Frontiers in Immunology*, vol. 14, Mar. 2023, https://doi.org/10.3389/fimmu.2023.1131051.

Link-Gelles, Ruth. “Interim Estimates of 2024–2025 COVID-19 ...” *MMWR. Morbidity and Mortality Weekly Report*, vol. 74, Feb. 2025, https://doi.org/10.15585/mmwr.mm7406a1.

Liu, Ge, Brandon Carter, and David K Gifford. “Predicted Cellular Immunity Population Coverage Gaps for SARS-CoV-2 Subunit Vaccines and Their Augmentation by Compact Peptide Sets.” *Cell Systems* 12, no. 1 (January 20, 2021): 102-107.e4. https://doi.org/10.1016/j.cels.2020.11.010.

Liu, Ge, et al. “Computationally Optimized SARS-CoV-2 MHC Class I and II Vaccine Formulations Predicted to Target Human Haplotype Distributions.” *Cell Systems*, vol. 11, no. 2, Aug. 2020, pp. 131-144.e6, https://doi.org/10.1016/j.cels.2020.06.009.

Lv, Hao, et al. “Application of Artificial Intelligence and Machine Learning for COVID-19 Drug Discovery and Vaccine Design.” *Briefings in Bioinformatics*, vol. 22, no. 6, Aug. 2021, https://doi.org/10.1093/bib/bbab320.

Malone, Brandon, et al. “Artificial Intelligence Predicts the Immunogenic Landscape of SARS-CoV-2 Leading to Universal Blueprints for Vaccine Designs.” *Scientific Reports*, vol. 10, no. 1, Dec. 2020, pp. 1–14, https://doi.org/10.1038/s41598-020-78758-5.

Martinez, Gustavo Sganzerla, et al. “PoxiPred: An Artificial-Intelligence-Based Method for the Prediction of Potential Antigens and Epitopes to Accelerate Vaccine Development Efforts against Poxviruses.” *Biology*, vol. 13, no. 2, Feb. 2024, https://doi.org/10.3390/biology13020125.

Matson, Robert P., et al. “A Deep Learning Approach Predicting the Activity of COVID-19 Therapeutics and Vaccines against Emerging Variants.” *Npj Systems Biology and Applications*, vol. 10, no. 1, Nov. 2024, pp. 1–10, https://doi.org/10.1038/s41540-024-00471-0.

Mirsanei, Zahra, Sima Habibi, Nasim Kheshtchin, Reza Mirzaei, Samaneh Arab, Bahareh Zand, Farhad Jadidi-Niaragh, et al. “Optimized Dose of Dendritic Cell-Based Vaccination in Experimental Model of Tumor Using Artificial Neural Network.” *Iranian Journal of Allergy, Asthma, and Immunology* 19, no. 2 (April 16, 2020): 172–82. https://doi.org/10.18502/ijaai.v19i2.2770.

Mørk, Sofie Kirial, et al. “Personalized Therapy with Peptide-Based Neoantigen Vaccine (EVX-01) Including a Novel Adjuvant, CAF®09b, in Patients with Metastatic Melanoma.” *OncoImmunology*, vol. 11, no. 1, Jan. 2022, https://doi.org/10.1080/2162402x.2021.2023255.

Mou, Lisha, et al. “Integrating Genomics and AI to Uncover Molecular Targets for mRNA Vaccine Development in Lupus Nephritis.” *Frontiers in Immunology*, vol. 15, Oct. 2024, https://doi.org/10.3389/fimmu.2024.1381445.

Meng, Lingxin, et al. “Individuals Carrying the HLA-B*15 Allele Exhibit Favorable Responses to COVID-19 Vaccines but Are More Susceptible to Omicron BA.5.2 and XBB.1.16

Infection." *Frontiers in Immunology*, vol. 15, Aug. 2024, https://doi.org/10.3389/fimmu.2024.1440819.

Mirsanei, Zahra, et al. "Iranian Journal of Allergy, Asthma and Immunology." *Iranian Journal of Allergy, Asthma and Immunology*, Apr. 2020, pp. 172–82, https://doi.org/10.18502/ijaai.v19i2.2770.

Montesi, Giorgio, et al. "Predicting Humoral Responses to Primary and Booster SARS-CoV-2 mRNA Vaccination in People Living with HIV: A Machine Learning Approach." *Journal of Translational Medicine*, vol. 22, no. 1, May 2024, pp. 1–9, https://doi.org/10.1186/s12967-024-05147-1.

Ong, Edison, et al. "COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning." *Frontiers in Immunology*, vol. 11, July 2020, https://doi.org/10.3389/fimmu.2020.01581.

"Predicting a Positive Antibody Response After 2 SARS-CoV-2... : Transplantation." *LWW*, https://journals.lww.com/transplantjournal/fulltext/2022/10000/Predicting_a_Positive_Antibody_Response_After_2.21.aspx. Accessed 26 Apr. 2025.

Roper, Willem. "Infographic: Vaccine Hesitancy Persists." *Statista*, 23 Feb. 2021, https://www.statista.com/chart/24263/coronavirus-vaccine-survey/.

Singh, Prateek, et al. "A Machine Learning-Based Approach to Determine Infection Status in Recipients of BBV152 (Covaxin) Whole-Virion Inactivated SARS-CoV-2 Vaccine for Serological Surveys." *Computers in Biology and Medicine*, vol. 146, July 2022, p. 105419, https://doi.org/10.1016/j.compbiomed.2022.105419.

Steinert, Janina I., et al. "COVID-19 Vaccine Hesitancy in Eight European Countries: Prevalence, Determinants, and Heterogeneity." Science Advances, Apr. 2022.

Tataje-Lavanda, Luis, et al. "Identification and Evaluation In-Vitro of Conserved Peptides with High Affinity to MHC-I as Potential Protective Epitopes for Newcastle Disease Virus Vaccines." *BMC Veterinary Research*, vol. 19, no. 1, Oct. 2023, pp. 1–7, https://doi.org/10.1186/s12917-023-03726-w.

Tu, Zhiwei, et al. "Helicobacter Pylori-Targeted AI-Driven Vaccines: A Paradigm Shift in Gastric Cancer Prevention." *Frontiers in Immunology*, vol. 15, Nov. 2024, https://doi.org/10.3389/fimmu.2024.1500921.

"Vaccination and Immunization Statistics." *UNICEF DATA*, 9 July 2019, https://data.unicef.org/topic/child-health/immunization/.

Wistuba-Hamprecht, Jacqueline, et al. "Machine Learning Prediction of Malaria Vaccine Efficacy Based on Antibody Profiles." *PLOS Computational Biology*, vol. 20, no. 6, June 2024, https://doi.org/10.1371/journal.pcbi.1012131.

Yang, Zikun, et al. "An in Silico Deep Learning Approach to Multi-Epitope Vaccine Design: A SARS-CoV-2 Case Study." *Scientific Reports*, vol. 11, no. 1, Feb. 2021, pp. 1–21, https://doi.org/10.1038/s41598-021-81749-9.

Sanctions, Strategy, and Survival: The Economic Impact of the Ukraine War on Kazakhstan–Russia Relations By Sophie Cao

Abstract

This paper examines the economic impact on Kazakhstan stemming from Russia's Invasion of Ukraine since February 2022. Kazakhstan benefited from being a trade intermediary with Russia, enabling the latter to circumvent Western sanctions. Kazakhstan has witnessed an increase in foreign direct investment (FDI) from Russia, and an exodus of both skilled Russian professionals, particularly in technology-intensive sectors, and Russian companies. These developments could explain why Kazakhstan's economy is growing despite the war in Ukraine, and why Kazakhstan's currency, the Kazakh tenge (₸), has recovered from the initial depreciation in the wake of the war. As the war progressed, Kazakhstan's strategic importance for Russia as a regional economic hub to circumvent sanctions has been increasing. Beyond reaping economic benefits, Kazakhstan has also strengthened diplomatic ties with Russia. These developments raise questions about the effectiveness of the Western sanctions on Russia and signal Central Asia's geopolitical orientation away from the West.

Introduction

The War in Ukraine has economic consequences across Eurasia. Kazakhstan stands out among Russia's neighbors due to its substantial trade volume, which has grown since the Invasion. Though many sanctions have been placed on Russia, its economy has not yet completely collapsed because Kazakhstan still remains one of Russia's key economic partners, even in war. In addition, companies have used Kazakhstan to access the Russian market, bypassing sanctions ("The Mysterious middlemen").

It would be a mistake to downplay the importance Kazakhstan has on Russia's economy. Unlike its smaller and less-developed Central Asian neighbors, Kazakhstan has more infrastructure and trade capacity, as well as closer economic interlink with Russia. For example, the 2014 Eurasian Economic Union (EEU) treaty signed by Kazakhstan, Russia, and Belarus demonstrated the close political and economic alignment between Russia and Kazakhstan. Currently, Kazakh politician Bakhytzhan Abdiruly Sagintayev chairs the EEU's Board of the Eurasian Economic Commission, continuing this tight geopolitical relationship. Moreover, the two countries engage in large volumes of trade: in 2021, the total volume of trade between Russia and Kazakhstan reached $25.6 billion. These intertwined political and economic ties highlight how Kazakhstan is indispensable to Russia.

Understanding Kazakhstan's role in mitigating Western sanctions for Russia has crucial implications for policymakers, economic researchers, and business strategists as it can answer critical questions about sanctions' effectiveness. Recognizing Kazakhstan's importance for Russia's post-Invasion economy, the West could build up stronger trade relations with Kazakhstan and its Central Asian neighbors, minimizing Central Asian countries' reliance on Russia, thus making Western sanctions more effective.

This paper divides Kazakhstan and Russia's trade and investment relations across three distinct periods: before the invasion, immediately after February 2022, and during the prolonged war period. It analyzes trade data from the National Bank of Kazakhstan, the World Bank, the International Monetary Fund (IMF), and Haver Analytics. It also investigated related news articles and scholarly sources. Employing these sources, this paper will construct a clear narrative on economic and political relationships between Russia and Kazakhstan, while also exploring how the War in Ukraine has reshaped Kazakhstan's economy and trade.

Pre-Invasion

Kazakhstan is a Central Asian country, with a small but growing economy dominated by oil and gas sectors. In 2018, its gross domestic product (GDP) was $179.3 billion. Then, it increased by 1.3 percent to $181.7 billion in 2019 but dropped by 5.8 percent to $171.1 billion in 2020 due to the Covid-19 pandemic. The economy expanded by 15.2 percent in 2021, largely due to the soaring energy prices, since global demand for oil increased "as mobility steadily returns to pre-pandemic levels, amplified by supply disruptions in the U.S., as well as the OPEC+ group's intention to keep a significant production capacity off the market," observed the World Bank ("Kazakhstan GDP 1990-2025;" 10). Consequently, its GDP increased to $197.1 billion in 2021 ("Kazakhstan GDP 1990-2025").

Kazakhstan is classified as an upper-middle-income country by the World Bank ("The World Bank in"). Kazakhstan's GDP per capita has been increasing for most of the last few years. In 2019, its GDP per capita was $9812.6 ("Kazakhstan GDP 1990-2025"). In 2020, Kazakhstan's economy was already facing headwinds from volatile oil prices, which, coupled with the global impact of the COVID-19 pandemic, caused its GDP per capita to drop by almost $700, a 7 percent decrease. It later quickly rebounded to $10,207.9 in 2021, a 12.6 percent increase, according to the IMF ("Republic of Kazakhstan;" "Kazakhstan GDP 1990-2025").

Kazakhstan's imports of goods were 21.5 percent of GDP in 2021. Its exports accounted for 33.8 percent of GDP in 2021. Its current account was $2.7 billion in 2021, which is 1.4 percent of GDP (The Kazakh National Bank *Balance of Payments*). Russia was one of its main trading partners. Kazakhstan imported 42.5 percent of its goods from Russia in 2021 (Haver Analytics). Russia's main exports to Kazakhstan include consumer goods, raw materials, and capital goods ("Product Exports"). Russia was the third largest exporting market of Kazakhstan in 2021, after China and Italy, accounting for 10.7 percent of Kazakhstan's total exports (Haver Analytics).

FDI has played an important role in Kazakhstan. FDI was 0.9 percent of Kazakhstan's GDP in 2021 (Haver Analytics). The two industries that receive the most FDI are mining and quarrying, and manufacturing. Within the first sector, extraction of crude petroleum and natural gas dominates, receiving $12 billion in FDI in 2018 and $12.2 billion in 2019. In 2020, however, the FDI dropped to $6.5 billion, likely due to the COVID-19 Pandemic. In the manufacturing sector, the manufacture of basic metals and fabricated metals, except machinery and equipment dominates as it received $2.7 billion in FDI in 2018, which increased by 1.5 percent to $2.8

billion in 2019. The FDI in this industry decreased by 10.2 percent to $2.5 billion in 2020, likely due to the COVID-19 Pandemic's global effects (The Kazakh National Bank *Gross inflow of direct investment*).

Kazakhstan's other FDI sectors are also notable. The country has seen an increase in FDI in its warehousing and support activities for the transportation sector. The FDI for that sector has doubled, jumping from $60.2 million to $120 million, between 2018 and 2019; however, there was a dip in 2020 to $92.6 million, a 22.8 percent decrease; and then a jump again in 2021 to $142 million, a 53.1 percent increase. Kazakhstan's information tech and other services industry, however, has seen a different story: it stayed around $40 million prior to 2021, which was when it more than doubled to $90.8 million (The Kazakh National Bank *Gross inflow of direct investment*).

Despite the amount of FDI Kazakhstan receives, it still has a relatively small economy, which means it draws less attention from many Western companies. Russia, however, does not. The amount of Russian FDI amounts to 25.5 percent of Kazakhstan's total FDI. Russia has invested heavily in Kazakhstan, spending between $1-1.5 billion each year from 2017 to 2020 and $1.9 billion in 2021 (Haver Analytics). This sudden increase in 2021, in the middle of a global pandemic, suggests that Russia may have been preparing closer economic ties with Kazakhstan before invading Ukraine.

In January 2022, Russian intervention allowed Kazakh President Kassym-Zhomart Tokayev to stay in power, warming relations between these two nations, demonstrating that influence over Kazakhstan is central to Russia's geopolitical calculus (Mallinson). Therefore, Russia and Kazakhstan have grown closer even before Russia was widely condemned for its invasion.

Immediately After Invasion

Kazakhstan was immediately impacted after Russia's War in Ukraine in February 2022, experiencing both adverse and beneficial effects. Following the Russian ruble (₽), the Kazakhstan tenge depreciated in February 2022. The exchange rate during January 2022 was at ₸435 to $1. One month later, the tenge depreciated, falling to ₸490 to $1. Overall, the tenge has oscillated below the pre-war levels ("Kazakhstani Tenge/U.S. Dollar"). The tenge's fluctuations reveal its economic vulnerability. The simultaneous deprecation of Russian ruble and the tenge discloses Kazakhstan's economic exposure and interdependence with Russia as the tenge mirrored the ruble's decline. Consequently, the Kazakh National Bank injected $138 million of its foreign reserves into the domestic market to stabilize and ease the tenge a few hours after the beginning of the invasion. Four days later, on February 28th, 2022, the Kazakh National Bank was prompted to inject another $98 million (Najibullah).

However, there are positive economic aftershocks as well. On the Russian side, thousands of companies or corporations, large or small, have flocked to Kazakhstan, to continue doing business while avoiding sanctions. In the three months after February 2022, the number of foreign companies in Kazakhstan grew from 22.4 thousand to 25 thousand, a 2.6 thousand

increase. In contrast, between January 2021 and February 2022, the number grew from 19.6 thousand to 22.4 thousand, a 2.8 thousand increase. While foreign companies encompass many nations, the increase comes mainly from Russian companies, which have reached 13 thousand in May 2022 ("Russian companies are actively").

Some Russian companies relocated to Kazakhstan to provide Russia with drones and microchips, which are needed to sustain its war effort while mitigating the effect of Western sanctions, indicating that trade with Kazakhstan allows Russia to mitigate the effects of sanctions. In 2021, $35 million of drone and microchip components were imported to Kazakhstan while $75 million was imported in 2022, a 214.3 percent increase, suggesting that Kazakhstan permits Russia to get its hands on sanctioned technologies ("Kazakhstan becomes pathway"). This has not been without consequences. Kazakh companies that have helped Russia evade sanctions have already been targeted by the United States Department of State ("New Measures to Degrade").

Trade in the telecommunications sector increased between the first quarter of 2021 and first quarter of 2022, when the War in Ukraine commenced. In the first quarter of 2021, Kazakhstan exported $8.8 million in the telecommunications sector to Russia; the amount rose to $9.5 million in the first quarter of 2022, a 7.9 percent increase (The Kazakh National Bank *External trade in services for 1-st quarter 2021-2022*). A similar trend can be seen in the trade with computers. During the first quarter of 2021, Kazakhstan exported $661.4 thousand to Russia, but the amount grew to $1.2 million in the first quarter of 2022, a 188 percent increase (The Kazakh National Bank *External trade in services for 1-st quarter 2021-2022*). The increase in technology exports to Russia demonstrates how Kazakhstan is a potential passage to avoid sanctions.

Prolonged War Period

During the entire 2022, Russia and Kazakhstan had $26 billion in trade, including exporting goods that could help Russia sustain its war effort, suggesting that Kazakhstan inadvertently aided Russia's war effort (Mallinson).

Since invading Ukraine, Russia's exports to Kazakhstan have decreased, which can be highlighted in the trade in telecommunications and computers. In the telecommunications sector during 2022, Kazakhstan imported $29.2 million from Russia (The Kazakh National Bank *External trade in services for 1-st quarter 2021-2022*). In 2023, it imported $25.5 million from Russia, a 12.7 percent decrease. This is a significant decrease compared to the $31.7 million imported in 2021 (The Kazakh National Bank *External trade in services for 2021-2022*). Albeit it imported $28.90 million from Russia in 2024, a 13.2 percent increase, there is an overall decrease since the beginning of the Invasion of Ukraine (The Kazakh National Bank *External trade in services for 2023-2024*). To keep Russia's technology sector afloat, Kazakhstan has imported Russian computers too. In 2023 and 2024, Kazakhstan imported around $60 million in this sector from Russia (The Kazakh National Bank *External trade in services for 2023-2024*). This is roughly a 30.5 percent decrease from 2022, when Kazakhstan imported $85.2 million in

computers from Russia (The Kazakh National Bank *External trade in services for 1-st quarter 2021-2022*). This decrease could be attributed to Russia's brain drain, with young, educated professionals, who could likely work in this sector, leaving Russia (Prokopenko). Another factor could be the fact that President Putin has prepared Russia's economy for total war, so almost all industries have been transformed to aid Russia's war effort.

In contrast, Kazakhstan exports to Russia have increased. The share of exports to Russia from Kazakhstan increased from 10.7 percent of Kazakhstan's total exports in 2021 to 12.7 percent in 2023 (Haver Analytics). In 2022, Kazakhstan exported $33.7 million to Russia in telecommunications. In 2023, Kazakhstan exported $33.5 million in the same sector to Russia; the amount stayed approximately the same. Yet, that amount has grown to $58.4 million in 2024, a 74.4 percent increase (The Kazakh National Bank *External trade in services for 2023-2024*). What is more striking is the computer sector. During the entire year of 2022, Kazakhstan exported $5.8 million to Russia (The Kazakh National Bank *External trade in services for 1-st quarter 2021-2022*). In 2023, Kazakhstan exported $41.2 million to Russia, which is a 714.1 percent increase. The amount has increased to $95.6 million in 2024, a 232.2 percent increase, demonstrating that Kazakhstan is becoming an ever-larger technology trading partner for Russia.

Looking beyond goods trade, compared with Russia's decreasing exports to Kazakhstan, Russia's FDI into Kazakhstan skyrocketed since the Invasion. In 2022, Kazakhstan received $1.5 billion in FDI from Russia, 24 percent of Kazakhstan's total FDI. Russia was Kazakhstan's fourth biggest investor. In 2023, Kazakhstan received $3.1 billion in FDI from Russia — accounting for 35.2 percent of total FDI — which is almost a 200 percent increase from the amount Russia invested in 2022. Russia has become Kazakhstan's second biggest foreign investor (Haver Analytics). In 2024, Kazakhstan received $4.5 billion in FDI from Russia, a 300 percent increase from 2022 (The Kazakh National Bank *Gross inflow*). Russia is now Kazakhstan's largest investor (Haver Analytics). The increase in Russian investment indicates that Kazakhstan is becoming more important economically and geopolitically for Russia: if it were not, then Russia would be investing considerably less money in Kazakhstan. In Kazakhstan's perspective, Russia has become a vital economic partner, further strengthening the ties between the two nations.

Kazakhstan's economy is benefitting from the exodus of Russian companies as technologies are being transferred to Kazakhstan. By March 2024, over $1.5 billion in investments, from 41 Russian companies, had been relocated to Kazakhstan. These firms are in a variety of different sectors, including manufacturing, IT, mining, and digital technology (Abbasova). Though some jobs are blue collar jobs, many in these sectors are white collar jobs, which are being added to Kazakhstan's job market. As a result, young, educated professionals not only bring their expertise to Kazakhstan but also add to Kazakhstan's human capital, thus benefiting it greatly.

Geopolitically, Russia has greater influence and leverage over Kazakhstan. Russia controls 25 percent of Kazakhstan's uranium production after President Tokayev allowed strategic licenses to be sold. Rosatom, the Russian state nuclear company, acquired a 49 percent

share in the company holding two licenses for the development of Kazakhstan's giant Budenovskoye uranium field at the end of 2022. In addition, President Tokayev has ceased to mention democracy in his speeches while creating an ever more authoritarian Kazakhstan by imposing stricter rules for unsanctioned assemblies (Mallinson). Therefore, Kazakhstan's political atmosphere is becoming more like President Putin's autocratic Russia. Even after the War in Ukraine has dragged on for over three years, the diplomatic relations between Russia and Kazakhstan have strengthened though Russia is isolated.

Data Analysis

As demonstrated above, Kazakhstan's economy has benefitted since Russia's War in Ukraine that commenced in February 2022. This relationship will be verified through data analysis in this section. In addition, the outcome of the data analysis can help explain how much of the overall economic landscape is influenced by the closer relationship since the Invasion.

The data used for this analysis span from 2005 through 2023, with certain series extending into early 2024. The selection of 2005 as a starting point reflects a period of relative macroeconomic stability for Kazakhstan, contrasting with the earlier volatile years, 1991-2004 ("GDP growth (annual %)"). This approach allows clearer analysis of the impact of the War in Ukraine beginning in February 2022.

A correlation matrix was employed to find which variables had strong relationships with each other. The variables used are GDP, in billions of US dollars; GDP per capita; GDP growth, in percentage; FDI from Russia; Kazakhstan's current account; and unemployment. A dummy variable called "Post 2022" takes on two values: 0, if the year is prior to 2022, and 1, if the year is 2022 or later. Additionally, it was introduced to explicitly account for the economic shock associated with the War in Ukraine, aligning with standard econometric practices.

First, the correlation analysis reveals a strong positive relationship between Russian FDI and Kazakhstan's GDP as $r = 0.776$, suggesting that higher Russian investments coincide with increased GDP levels. Second, a 0.799 value between FDI from Russia and the year after 2022 demonstrates a strong positive relationship between the year being closer to present day and the amount of FDI from Russia. However, not all values are statistically significant. Since the current account's coefficient is 0.030 when correlated with Kazakhstan's GDP, it suggests that the current account is an insignificant predictor in this model. The -0.835 value between unemployment and GDP demonstrates a strong negative correlation: the more unemployment, the less GDP.

A linear regression model was conducted with Kazakhstan's annual GDP growth set as the dependent variable; FDI from Russia; Kazakhstan's current account; and unemployment were set as the independent variable to capture the war's effect. Like the correlation analysis, the dummy variable called "Post 2022" was also used. It takes on two values: 0, if the year is prior to 2022, and 1, if the year is 2022 or later.

The regression model explains approximately 58.8 percent ($R^2 = 0.588$) of annual GDP growth variance. Thus, these four predictors — FDI from Russia, unemployment, the years being

after 2022, and Kazakhstan's trade balance — have moderate explanatory power. However, the regression model also highlights the presence of important yet unexplained variance, accounting for the remaining 41.2 percent, possibly due to unobserved factors or omitted variables.

First, the FDI coefficient is 0.00228. Thus, on average, each additional $1 million of Russian FDI is associated with a 0.00228 percentage-point increase in GDP growth — a statistically non-significant effect (p = 0.132). Second, the Post 2022 dummy coefficient is -1.276, which suggests that, after 2022, GDP growth was 1.28 percentage points lower on average, but this effect is also nonsignificant. Notably, unemployment's positive coefficient 2.351 is paradoxical and counterintuitive as it suggests that for each unit of unemployment, Kazakhstan's GDP will grow by 2.351 percent: the greater the unemployment, the greater the GDP growth. This result contradicts theoretical expectations, as well as the correlation's findings, suggesting potential omitted variables or complex underlying dynamics not captured by the simple linear regression model.

Correlation Matrix

		GDP (in $ Billion)	GDP Growth (annual %)	GDP Per Capita (in $)	FDI from Russia	Year	Current Account	Unemployment
GDP (in $ Billion)	Pearson's r	—						
	df	—						
	p-value	—						
GDP Growth (annual %)	Pearson's r	−0.338	—					
	df	17	—					
	p-value	0.156	—					
GDP Per Capita (in $)	Pearson's r	0.974 ***	−0.274	—				
	df	17	17	—				
	p-value	<.001	0.256	—				
FDI from Russia	Pearson's r	0.776 ***	−0.172	0.667 **	—			
	df	17	17	17	—			
	p-value	<.001	0.482	0.002	—			
Year	Pearson's r	0.741 ***	−0.569 *	0.583 **	0.799 ***	—		
	df	17	17	17	18	—		
	p-value	<.001	0.011	0.009	<.001	—		
Current Account	Pearson's r	0.030	0.438	0.147	−0.211	−0.365	—	
	df	17	17	17	18	18	—	
	p-value	0.902	0.061	0.549	0.372	0.114	—	
Unemployment	Pearson's r	−0.835 ***	0.628 **	−0.782 ***	−0.583 **	−0.828 ***	0.193	—
	df	17	17	17	18	18	18	—
	p-value	<.001	0.004	<.001	0.007	<.001	0.416	—

Note. * p < .05, ** p < .01, *** p < .001

Table 1: Correlation matrix of the employed model

Conclusion and Recommendations

Kazakhstan and Russia's close economic relationship started before the War in Ukraine and continue to become more mutually dependent. Russia has been investing more money into Kazakhstan after 2022 while trade between these two nations has continued to surge, in particular, exports from Kazakhstan to Russia. Though the tenge depreciated immediately after February 2022, when the full-on Invasion in Ukraine was announced, Kazakhstan's economy has greatly benefitted, nonetheless. Many young, educated Russians have fled to Kazakhstan, transferring in technical expertise and human capital there. Additionally, Russian companies

have moved to Kazakhstan, creating new jobs for Kazakhstan's economy.

The correlation matrix has found that FDI from Russia, whether before or after invasion, and Kazakhstan's GDP have a strong positive relationship. The closer it is to the present day, the more FDI from Russia. Greater FDI from Russia correlates to Kazakhstan's GDP increasing. Yet, the linear regression has found that greater unemployment would statistically lead to more GDP growth for Kazakhstan. Thus, future studies are recommended to model the linear regression, perhaps with more data or a different analysis, to either solve or explain the paradox between unemployment and GDP growth. In this linear regression model, unemployment emerges as the statistically robust predictor of Kazakhstan's GDP. In addition, it is recommended to either wait a few years and then collect the new data or zoom into a specific year and use only data from that year. A new linear regression can be modeled with the new data to see if it is possible to find any new, interesting, and important findings. While the correlation graph can tell us that Russian FDI and Kazakhstan's GDP growth have a strong relationship, it is still important to find causational relationships as well. Likewise, as the year is closer to present day, the correlation graph indicates that Russia invests more into Kazakhstan.

Nevertheless, it is important to understand whether the FDI is purely for economic purposes or there are other political, social, or less well-known intentions. Diplomatically, Russia and Kazakhstan have grown closer. Understanding the importance of Kazakhstan to Russia and Kazakhstan's own economy, politics, and culture allows not only businesses but also policy makers to cultivate economic and geopolitical ties with it. Russia's relationship with Kazakhstan demonstrates that cultivating these ties requires both time and effort. In conclusion, examining these economic relations between Russia and Kazakhstan can help researchers identify shortcomings of current sanctions, as well as discovering new strategies to enhance the effectiveness of sanctions.

Works Cited

Abbasova, Vusala. “Russian Business Presence in Kazakhstan Drops to Six-Year Low.” Caspian News, 16 Apr. 2025, https://caspiannews.com/news-detail/russian-business-presence-in-kazakhstan-drops-to-six-year-low-2025-4-16-32/. Accessed 20 May 2025.

“GDP Growth (Annual %) - Kazakhstan | Data.” World Bank, https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=KZ. Accessed 20 May 2025.

Haver Analytics. Haver Analytics.

Kazakh National Bank (2024). Balance of Payments of the Republic of Kazakhstan. [Data set]. https://nationalbank.kz/en/news/platezhnyy-balans-vn-sektora. Accessed 20 May 2025.

Kazakh National Bank (2022). External trade in services of the Republic of Kazakhstan by types of services for the following countries for 1-st quarter 2021-2022. [Data set]. https://nationalbank.kz/en/news/platezhnyy-balans-vn-sektora.

Kazakh National Bank (2023). External trade in services of the Republic of Kazakhstan by types of services for the following countries for 2021-2022. [Data set]. https://nationalbank.kz/en/news/platezhnyy-balans-vn-sektora.

Kazakh National Bank (2024). External trade in services of the Republic of Kazakhstan by types of services for the following countries for 2023-2024. [Data set]. https://nationalbank.kz/en/news/platezhnyy-balans-vn-sektora.

Kazakh National Bank (2025). Gross inflow of direct investment in Kazakhstan from foreign direct investors. [Data set]. https://nationalbank.kz/en/news/pryamye-investicii-po-napravleniyu-vlozheniya. Accessed 21 May 2025.

“Kazakhstan Becomes Pathway for Supply of Russia’s Defence Sector with Drones & Microelectronics despite Western Sanctions, New Investigation Shows.” Business & Human Rights Resource Centre, 19 May 2023, https://www.business-humanrights.org/en/latest-news/kazakhstan-becomes-pathway-for-supply-of-russias-defense-sector-with-drones-microelectronics-despite-western-sanctions-new-investigation-shows/. Accessed 20 May 2025.

“Kazakhstan GDP.” MacroTrends, https://www.macrotrends.net/global-metrics/countries/KAZ/kazakhstan/gdp-gross-domestic-product. Accessed 20 May 2025.

“Kazakhstan GDP 1990-2025.” MacroTrends, https://www.macrotrends.net/global-metrics/countries/KAZ/kazakhstan/gdp-gross-domestic-product. Accessed 20 May 2025.

“Kazakhstan GNI 1993-2025.” MacroTrends, https://www.macrotrends.net/global-metrics/countries/kaz/kazakhstan/gni-gross-national-income. Accessed 20 May 2025.

"Kazakhstan Interest Rate." Trading Economics, https://tradingeconomics.com/kazakhstan/interest-rate. Accessed 20 May 2025.
"Kazakhstani Tenge/U.S. Dollar Advanced Charting | KZT to USD." Wall Street Journal, n.d., https://www.wsj.com/market-data/quotes/fx/KZTUSD/advanced-chart. Accessed 20 May 2025.
Mallinson, Kate. "Russia's Influence in Kazakhstan Is Increasing despite the War in Ukraine." Chatham House, 8 Dec. 2022, https://www.chathamhouse.org/2024/02/russias-influence-kazakhstan-increasing-despite-war-ukraine.
Najibullah, Farangis. "Central Asia Takes Economic Hit From Russian War In Ukraine Sooner Than Expected." Radio Free Europe Radio Free Library, 1 Mar. 2022, https://www.rferl.org/a/central-asia-migrants-ruble-impact/31730968.html. Accessed 20 May 2025.
"New Measures to Degrade Russia's Wartime Economy." United States Department of State, https://2021-2025.state.gov/new-measures-to-degrade-russias-wartime-economy/. Accessed 20 May 2025.
"Product Exports to Kazakhstan 2022." WITS Data, https://wits.worldbank.org/CountryProfile/en/Country/WLD/Year/2022/TradeFlow/Export/Partner/KAZ/Product/all-groups. Accessed 21 May 2025.
Prokopenko, Alexandra. "The Great Russian Brain Drain." George W. Bush Presidential Center, 2024, https://www.bushcenter.org/catalyst/the-great-gray-wave/the-great-russian-brain-drain/. Accessed 12 Jun. 2025.
"Republic of Kazakhstan and the IMF." International Monetary Fund, https://www.imf.org/en/Countries/KAZ#countrydata. Accessed 20 May 2025.
"Russia Product Exports and Imports to Kazakhstan 2021." WITS Data, https://wits.worldbank.org/CountryProfile/en/Country/RUSSIA/Year/2021/TradeFlow/EXPIMP/Partner/KAZ/Product/All-Groups. Accessed 7 Aug. 2025.
"Russian companies are actively migrating to Kazakhstan." Kapital.Kz, https://kapital.kz/business/106788/rossiyskiye-kompanii-aktivno-migriruyut-v-kazakhstan.html. Accessed 20 May 2025.
"The Mysterious Middlemen Helping Russia's War Machine." The Economist, 19 Aug. 2024, https://www.economist.com/europe/2024/08/19/the-mysterious-middlemen-helping-russias-war-machine. Accessed 20 May 2025.
"The World Bank in Kazakhstan." World Bank, https://www.worldbank.org/en/country/kazakhstan/overview. Accessed 8 Aug. 2025.

Analyzing New Jersey's Emission Reduction Goals by Combining Heterogeneous Data
By Neil Mehra

Abstract

Greenhouse gas emissions are the largest source of anthropogenic climate change, which has contributed to temperatures rising over 1.5C globally since the start of the industrial age. Such rising temperatures have significant long-term health and economic consequences, among them increased risk of death from heatstroke and loss of arable land through shifting biomes. To combat this, frameworks such as the Paris Agreement have been engineered to halt the rising temperatures, with many countries taking part in action to reach net-zero carbon emission status. In 2006 in New Jersey, the 80x50 framework was established with the goal to reduce net carbon emissions by 80% by 2050, a goal believed to align with the Paris Agreement. In 2020, the 80x50 framework was revised with a new set of proposed policies and milestones that kept the same 80% reduction goals in mind. Each emissions sector (commercial/residential, power generation, transportation, etc) has a set of individuals and milestones by certain years, with the earliest being 2025 and latest being 2050. This paper seeks to investigate the transportation and the residential/commercial sectors and their progress over the last 20 years to determine whether the 80x50 milestones are on target.

Introduction

The industrial age is estimated to have begun in the mid 18th century and ended in the late 20th century. This period of human history was characterized by a shift from agriculture-based economies to a more machinated society with a versatile output of goods. The industrial revolution not only massively changed the production sector by enabling the mass production of various products, but it also allowed for development of superior technology almost universally (van Neuss, 2015, pp. 3-6). The rapid expansion of the job market due to factory vacancies drove the development of densely-populated urban zones, a movement which also furthered the need for improved transcontinental and international transport. This was seen in the advancements made in railroad infrastructure, maritime trade through the creation of modern cargo ships. and eventually in the aviation industry (Alvarez-Palau, 2020). In addition, the growth of factories also led to the increased amount of local transport, namely in the form of automobiles. Even in agriculture, an industry which was heavily reduced during this time, machination and pest management paved the way for large increases in the efficiency of food production (OER Commons, 2018). All of these changes in infrastructure coupled with revolutions in healthcare, education, and communication have led to an overall increase in the quality of life (Cutler & Miller, 2005).

That being said, there are significant drawbacks that are also associated with the industrial revolution (Obschonka, 2018). Despite the many successes this period had, they also somewhat contributed to massive problems which have developed into major global issues today. The prime example of this is the constantly rising average global temperatures since the

start of the industrial age. As of 2024, the average global surface temperature is about 1.55 degrees Celsius higher than its counterpart in 1850 (Barrage & Nordhaus, 2023). While a certain amount of temperature variation as a result of changing long-term climates is somewhat expected, such a drastic increase in less than two centuries dismisses this idea. The rising of surface temperatures has been attributed both directly and indirectly to anthropogenic causes (US EPA, 2025). The earth's surface temperature maintains its warmth via the greenhouse effect. This effect stipulates that certain gaseous compounds in the atmosphere act as a blanket to trap heat near the surface of the earth. While this effect may often receive a negative connotation due to its association with anthropogenic climate change, it is essential to human life as without the presence of these gasses the earth would not retain enough heat to sustain life. However, human action during the industrial age has amplified the effects to the point that they are severely harmful to human life (US EPA, 2025).

Impacts to Humans

The impacts to human life that result from rising surface temperatures range from direct illnesses to indirect social costs through the damaging of infrastructure. Increasing temperatures globally have caused shifts in typically tropical diseases to spread further from the equator to more temperate climates, namely Malaria (Rocque et al., 2021). In already temperate/cooler climates such as that of NJ (the focus of this paper), heat-related illnesses such as heatstroke have become significantly more common and threaten populations to a larger extent, most notably the elderly. New Jersey has issued official warnings to the elderly and the demand for better urban cooling in these areas has increased drastically. In addition to these, there are also more indirect health concerns posited by the overall increase in surface temperatures, including more favorable conditions for harmful bacteria, vector species, asthma, toxic algal blooms, and respiratory diseases such as asthma (Rocque et al., 2021). There are also environmental costs to consider when evaluating the total impact of rising surface temperatures on humans.

As temperatures increase across the earth, biomes (areas characterized by particular longlasting precipitation and temperature) shift. Generally, temperature and tropical zones have expanded farther from the equator to replace cooler biomes, such as the northern boreal forests and snowy tundras in Canada (Bonannella et al., 2023). These changing biomes lead to changing habitats for all species living in these areas. While these conditions may benefit or harm certain species, general trends point to most experiencing the latter (Bonannella et al., 2023). A primary example of this in North America is the rapidly diminishing population of the monarch butterfly due to habitat loss of both breeding and feeding habitat as a result of climate change. Shifts in biomes also impact agricultural output, as the efficiency of certain crops often decrease as areas may shift away from ranges of tolerance (Bonannella et al., 2023). These impacts have social and economic implications. According to projections, each metric ton of CO2e emitted costs about $258. In the United States alone, emissions reach over 6 Billion MT annually. This equates to about 1.6 Trillion USD annually in social costs (Barrage & Nordhaus, 2023), although this is a global estimate and regional differences may decrease this number. More importantly, climate

change related deaths are projected to reach 83,000,000 by the year 2100. This number could be as high as 275 Million or as low as 20 million depending on what is done to combat climate change in the coming years (Barrage & Nordhaus, 2023).

Causes of Global Warming

Global warming and the magnification of the greenhouse effect are caused by greenhouse gasses (GHGs) which are quantified on the basis of their potency and how long they remain in the atmosphere. The basic unit for quantifying GHG emission impact is via CO2e or Carbon Dioxide equivalent. One metric ton of CO2e equates to how much global warming potential one metric ton of CO2 has over the span of 100 years. Other gasses use CO2e as a scale to determine their potency and longevity in the atmosphere. For example, one metric ton of methane (CH4), one of the most prominent GHGs, has the same polluting power of 25 metric tons of CO2e. This means that even though methane takes up a smaller share of the atmosphere compared to Carbon Dioxide, it still plays a significant role in the greenhouse gas effect due to its potency.

The sources of greenhouse gasses are numerous but can vary in relative presence depending on each individual area. Generally, the main sources for anthropogenic greenhouse gas emission are transportation, energy production, commercial/residential, industrial, landfills, black carbon, and agriculture (Barr et al., 2020). Each individual source has its own subcategories which consist of different GHG sources, but they in total contribute to over 99% of total emissions. For example, agricultural emissions are most commonly related to land-use and the majority of CO2e is produced via methane. Alternatively, the energy and transportation sectors (which are most commonly the largest shareholders of global CO2e emissions), pollute primarily via CO2 itself. Each of these sectors has, since the start of the industrial revolution, grown extensively in production, mechanization, and efficiency. However, overall sustainability was not as significant of a theme during this period, leading to drastic increases in emissions overall (Graven et al., 2020). To put this into context, CO2 emissions alone in 1850 and 1900 were roughly 200 million and 2 billion metric tons per year. Today that number is close to 40 Billion. These numbers continue to follow positive trends although at less steep rates (Ritchie & Roser, 2020). Part of this is due to the general idea that populations and the global economy increasing are directly proportional to the increasing amounts of CO2e being emitted.

Legal Frameworks and the 80x50 Report

As a result of carbon emissions and subsequent climate change, legal frameworks both nationally and internationally have been implemented. The most popular of such frameworks is the Paris agreement which stipulates that countries should continuously slash emissions in an effort to become net 0 in the near future. The objective of these frameworks is to structure the development of smaller goals and policies. A more local example, particularly in select parts of the USA, is the idea of the 80x50 initiative (Barr et al., 2020). This stands for 80% reduction in carbon emissions by the year 2050. One state to implement this is New Jersey, a state known for

already being a leader in climate-friendly policies. Given its small size but large population density, New Jersey is unique in its approach to 80x50, having ambitious milestones that include having only EVs on the road by 2050, fully decarbonizing buildings, and slashing total emissions from 124 million metric tonnes (MMT) in 2006 to about 24 MMT in 2050 (Barr et al., 2020).

Gap and Justification of Gap

Given the perceived ambitiousness of New Jersey's 80x50 report, there are concerns that the milestones set by the report on the road to 2050 will not be reached in time. Currently, no cross-examination of information on New Jersey's carbon emission reduction efforts exists. This may prove detrimental as pivoting to a different strategy in the recent future would be more plausible if the goals themselves were found to not be as effective as what was previously intended. The report itself states that no large scale analysis has been done to determine whether or not these milestones are set to be achieved (Barr et al., 2020). However, data for multiple variables strongly correlated with carbon emission reduction is available on the internet for public use (Ritchie & Roser, 2020). This research paper intends to expand upon this gap by analyzing whether or not individual sectors that emit CO2e gasses are on track to reach their eventual targets by 2050. As of 2024, NJ has dropped from roughly 124 MMT of CO2e in 2006 to about 96 MMT. However, almost all of this progress had been made in the energy production sector as coal-power plants shifted to fossil fuels (and nuclear to an extent (USEIA, n.d.)). The other sectors have seen marginal decreases or have remained stagnant during this time. Thus, it is also vital to understand why the goals may not be reached if they are not projected to do so. To be more specific, this research paper will focus specifically on the transportation sector and the residential/commercial sectors as these are the largest contributors to carbon emissions in NJ (Barr et al., 2020). Data for other sectors exists but is not plentiful enough to conduct meaningful analysis on at this time.

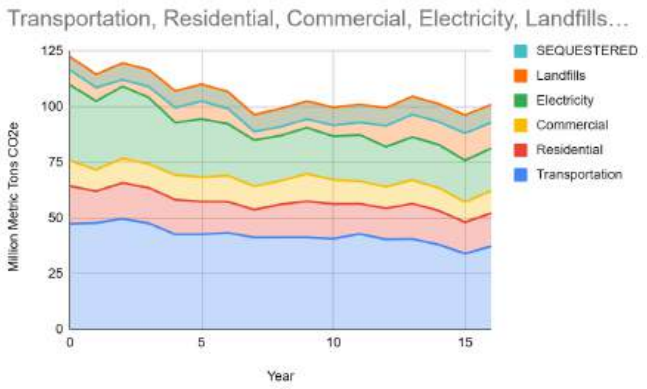


Fig 1: Layered map of CO2e emissions MMT by sector since 2006

Methodology

This study aims to analyze and quantify the progress that New Jersey has made in the past two decades in regards to the milestones outlined in the Global Warming Response Act 80x50. The objective will be to determine whether the milestones of the report are in line with the progress being made and the data produced and analyzed will be used to further explain the trends found. This study will specifically pertain to the transportation sector and the commercial/residential sectors (this was done due to the fact they have very similar patterns as will be highlighted). When performing a study such as this, comprehensive data is required in order to base judgements that can be used to derive meaningful conclusions (Wang et al., 2023). A sample size too small such as one that only incorporates a timeframe of 2-3 years or a limited amount of considered variables would lead to likely inaccuracies in analysis. Additionally, only drawing from one or two datasets would increase the likelihood of bias and would not derive any new meaning from the data (Wang et al., 2023). Due to the scale of the data analyzed being on a state level, it would only be feasible to draw data from sources already available publicly as opposed to physical experimentation. This meant that a meta-analysis of data was the most suitable methodology for conducting this paper. That being said, a traditional meta-analysis approach would have to be modified given the heterogeneity of the data being considered. Conventional meta-analyses of data involve searching multiple related studies for common trends or themes in conclusions found. However, the data used for this experiment had to be more widespread and thus it was more difficult to get them in similar formats to compare them. An approach that was considered was to calculate an I2 value which measures the exact amount of heterogeneity in the data (Ruppar, 2020). This would allow for direct comparison of the datasets without having to combine them. This was ultimately rejected because the data was too heterogenous and the common year variable made it significantly more justifiable to directly combine the datasets. While an I2 statistic is commonly used to quantify the heterogeneity of data in meta-analysis, it is ultimately inappropriate for this study as the datasets used vary significantly in format, scale, and measured variables. All datasets did however align along one temporal variable: year. This consistent annual categorisation across datasets allows for accurate chronological alignment which is a better solution than trying to statistically harmonise inherently incompatible formats. Ultimately, these factors led to the selection of a meta-analysis of heterogeneous data online by first collecting, cleaning, aggregating, combining, and then correlating and regressing the resulting datasets. In addition, alternative charts were also created to help explain the trends shown in the regression by relating it to other variables that were part of the data.

Methodology Justification

The proposed methodology of a meta-analysis consisting of a combination of largely heterogeneous data draws inspiration from similar studies conducted in the field of carbon emission analysis. The idea to combine data into a structured dataset before analyzing has proven to be significantly more accurate than analyzing individual datasets and creating predictive models from them (Nguyen et al., 2023). They conclude that models such as linear

regression which are used in this methodology gain significantly more accuracy when fed more comprehensive data from a larger amount of sources. Another key aspect of this paper is properly addressing the heterogeneity, or the fundamental differences in data structure, of the various files used for this project. Unlike typical meta-analyses which use meta-regression plots to evaluate the differences in data based on select columns, such shared columns are not present in this set of data due in part to the lack of long term publicly available data. Thus, multivariable regression will be used to create models for the net carbon emissions per sector and then will be used to highlight future trends. This was used in a similar paper by (Zhou et al., 2021), which successfully used regression analysis as a tool for finding the driving forces for rising CO2 emissions in China. The data itself will be compiled partially in excel and in Python using the pandas library. The pandas library of code allows for the importation of csv files which are transformed into DataFrame objects. These DataFrame objects can be used to perform concatenation of multiple datasets and then can be used in conjunction with other libraries to perform regression on specific columns (Singh et al., 2024).

Methodology Enactment

The first major step in enacting the methodology was to select parameters for the types of data that would be gathered. The source, structure, and content of the data was considered in order to select useful data which would not obscure results. This paper only focuses on transportation and residential/commercial sectors which outlined multiple key areas of interest. Only structured data was widely available online and was almost exclusively found in CSV file format. The small amount of data provided in the 80x50 report through tables gives the amount of CO2e per source for each sector in every year since 2006. This table was converted to a CSV file and served as the primary dataset from which the rest of the data was built. Another parameter is that each subsequent file found had to have some sort of annual categorization of data. For example, one dataset found from the afv (alternative fuel vehicle) report categorized the amount of electric vehicles purchased per year over the course of the period 2019-2024. Although not a broad enough of a time span to be considered ideal, electric vehicle growth before this period was comparatively negligible and thus such a time span was appropriate. Although 2015 saw slight increases in EV investment on both the east and west coast, the number of registered vehicles is only about 1% of the total on the road today according to the report.

In addition, data that provided itemized lists of certain information (such as the complete vehicle registration registry of New Jersey and the solar panel installation registry of New Jersey) had to be aggregated by the year. For example, each row that represented an individual car registration had to be combined. This structured the data so that each row contained a year and all of the aggregated statistics, including the total number of cars sold, the total amount of new cars registered, the amount of each type of vehicle being sold (EV or gasoline powered), and so forth. Datasets that were considered to not be from a trustworthy source or had too small of a sample size were rejected to avoid considerable bias. The primary search terms used were

“New Jersey Vehicle Data”, “New Jersey emissions data”, and “New Jersey space-heating and natural gas data”. Data that was not specifically restricted to New Jersey was also rejected because the proportion of emissions per sector is different compared to neighboring states. In total, 8 sources of data were used which varied in size from 20 rows (emissions by sector by year) to roughly 830000 (total vehicle registrations). The main sources of data were the NJDEP, SustainableJC, the afv report, and the 80x50 report itself.

The next step was to begin combining data into comprehensive datasets that could then be analyzed using regression. Two culminating datasets were produced to represent the transportation sector and the residential/commercial sectors respectively. Virtually all of both residential and commercial sector emissions are traced to natural gas being used for space heating and cooling so they were combined. The data had to be aggregated by year for both but more columns were added to avoid losing important information, such as counters to sum up the amount of vehicles that were unregistered and abandoned per year based on the discrepancy between registrations and total cars on the road. The transportation dataset was also used to calculate an experimental percentage of the amount of vehicles on the road and purchased per year that were EVs. The commercial/residential sectors were related to the expansion of solar panels as a clean energy investment. Solar panels are the primary source of clean energy on a domestic level (excluding nuclear power plants) and thus they stand the most significant chance of replacing natural gas as the source of energy for these sectors (The Conference Board, 2024).

A conversion factor of 0.375 kg of CO2e emissions per kilowatt hour of natural gas related energy production was found (Climeco). This coefficient was used in tandem with the wattage capability of existing solar panel installations to determine what level of investment in solar energy would be required to offset both sectors emissions by the year 2050. After this step, all 800000+ rows of data were condensed, combined, somewhat modified with conversion factors, and aggregated into more easily accessible data that could then be analyzed.

As stated before, the chosen method for analysis was regression. Before conducting the regression, several other charts were created to visualize the newly formed datasets and provide context for any conclusions that could be drawn from the regressions themselves. Such charts included an area graph that represented the proportions of electric vehicles and gas-powered vehicles as positively adding to the total number of vehicles whilst the amount of obsolete vehicles is also plotted on the negative y-axis.

Lastly, regression models were generated using the combined datasets. Regression was first used to calculate the predicted amount of CO2e emissions per year for each sector until 2050, and the final values were printed. This was accomplished by importing the complete csv files into a Google Colab code notebook and converting it to a DataFrame with the pandas library. Then with the addition of other libraries for model creation and visualization (numpy, matplotlib.pyplot, sklearn), these models were created.

Results

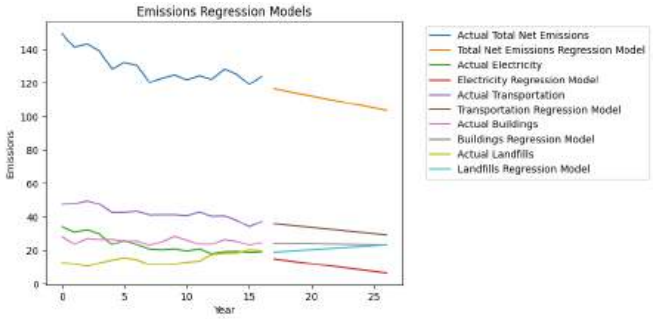


Figure 2: The overall regression model that plotted each individual sector of carbon emissions as a linear function is shown below:

This model shows overall decreasing linear trends in each sector as well as the net emissions, with the exception being landfills. The years are since 2006 which is the start of the available data for this model. This can be attributed to the growing problem of waste management and the subsequent methane emissions that follow as a result. The other sectors were hypothesized to have linearly decreasing trends, although the rate of decrease over time was less clear. The 80x50 report approximates the trend of carbon emissions to be decreasing and linear, although behavioral shifts and policies will create lag. That being said, time-series models carry the logical limitation that any future policy revisions or global irregularities such as the coronavirus pandemic cannot be predicted and thus are better taken as trends than predictive of exact values (Zillz, n.d.). Commercial and residential emissions in this scenario are aggregated under "Buildings" and it is apparent that the slope of this model is very minimal implying that there could be little to no change in this sector whereas the 80x50 report expects it to be a 0 emission sector by 2050 (Barr et al., 2020).

```
Predicted Total Net Emissions for 2045: [76.72107843]
Predicted Electricity Emissions for 2045: [-11.19215686]
Predicted Transportation Emissions for 2045: [15.56372549]
Predicted Buildings Emissions for 2045: [20.64681373]
Predicted Landfills Emissions for 2045: [32.23480392]
```

Figure 3: Predicted values for 2045 by the model in Figure 2

These are the predicted values of the regression model in Figure 2. The most notable observations include a very marginal drop in building emissions and a moderate (and still below expectation) decrease in both transportation and total net emissions.

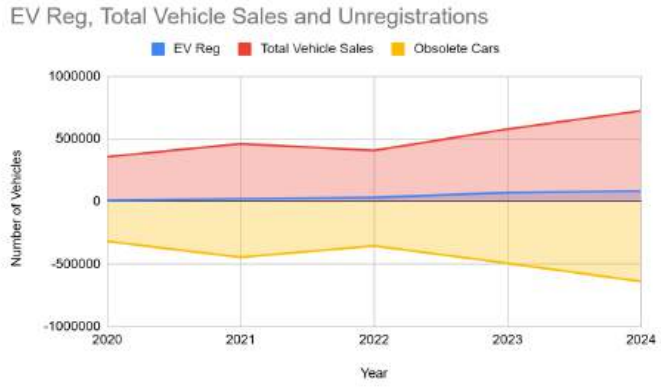


Figure 4: Area graph of EV and Vehicle sales over time.

Figure 4 illustrates how over time, even though the market share of electric vehicles in comparison to gas powered vehicle sales has increased from about 2% to 11% of all sold vehicles since 2020 (Barr et al., 2020), the total number of vehicle sales has also decreased and thus minimized the positive impacts that would have been associated with lesser gas powered vehicles on the road.

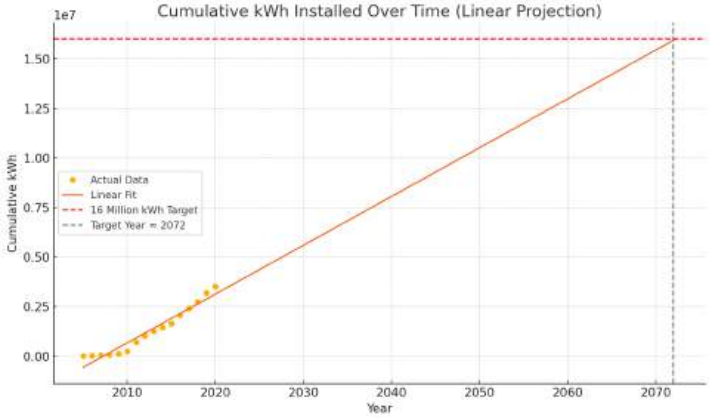


Figure 5: Linear regression of solar panel installations

This linear regression model shows the total solar panel wattage installed since 2005 and demonstrates a linear line of best fit. A linear line was chosen as opposed to an exponential one due to the amount of new solar panels every year stabilizing after having shot up massively since 2022 as evidenced by the chart. Using conversion factors, it is estimated that each kW of wattage saves 0.816 tons of metric CO2 per year. Given the 24.6 MMT of residential and commercial sector emissions to date, the amount of solar installations required to offset this stands at around 16 million kW. The purple vertical line represents the year at which there would be enough solar panel installations to completely shift from natural gas production in buildings to pure electrification and solar panels (When the total available kWh in NJ reaches 16 million). This is a theoretical estimate that assumes ideal conditions where solar wattage is directly used to offset natural gas emissions in the residential and commercial sectors. In reality, grid

transmission losses, electrification of appliances, and peak usage misalignment could increase or decrease this number. This aligns with the 80x50 goal to adopt clean energy in this sector, but is significantly slower than anticipated with the predicted year of reaching net 0 emissions being the year 2072.

Conclusion

To answer the initial research question regarding whether or not 80x50 goals are on target to be met in the transportation and building sectors, it is ultimately not likely given the current trends seen. New Jersey was predicted to reach 24.1 MMT of CO2e emitted by the year 2050 (Barr et al., 2020) but these regression models suggest that number to be over triple at about 76 MMT. The transportation sector is expected to see significant progress in the future due to the further incentivization of EVs as proposed by the 80x50 report which suggests that 85% of vehicles sold by the year 2030 will be electric vehicles (Barr et al., 2020). That being said, it is extremely unlikely for this to occur given the data collected as 2024 estimates place that number at around 11% instead. Meanwhile, the building sector remains largely untouched and this is likely to remain a significant obstacle in achieving 80% reduction in carbon emissions by 2050. New Jersey as a state has an infrastructure heavily dependent on natural gas from the past, and electrifying buildings to be net neutral entails a heavy cost. It is also important to note that even if buildings are electrified, unless their new energy source is renewable and a non-pollutant, their emissions would simply be transferred to the energy production sector.

The main questions that arise from these conclusions pertain to how the situation could be remedied without calling for a complete overhaul of policies. It is also crucial to understand why there have been significant difficulties when tackling this problem. This research implies through the chart and data collected that the battle with transportation emissions is less intertwined with electric vehicles than had previously been thought. While electric vehicles do not produce as much CO2e as their gas-powered counterparts, factors such as car prices, charging times, and location of EV chargers serve as heavy disincentives. The area chart also implies that the total share of vehicles is increasing, presumably due to a steadily increasing population. What this means is that the total number of vehicles must come down in order to reduce the net carbon emissions in this sector. There have been multiple proposed solutions for this endeavor, most notably the ideas of a 10 minute walkable city, increased work-from-home opportunities, and increased investment into public transport (J. Shope, personal communication, March 15, 2025). All of these would decrease the total amount of vehicles on the road, decrease the length of still existing vehicle trips, and encourage shifts to more environmentally-friendly transport, thus allowing total emissions to drop. On the building's front, both regression models imply that even if electrification of the grid is occurring, the rate at which it does occur is slow enough to not reach the 80x50 goals. The implications of this are that more investment is required in older buildings to convert from natural gas dependent temperature control to cleaner energy.

This research also supports the idea that the reason why NJ appears so off target may not

have to do with the goals themselves but rather the awareness of their existence in the public (J. Shope, personal communication, March 15, 2025). As part of constructing this paper, Professor James Shope, who worked on the 80x50 report and a subsequent evaluation of NJ's emission reduction goals, was interviewed. He expressed the idea that one of the key reasons why NJ appears stuck is the fact that the general public aren't really aware of how they could help (J. Shope, personal communication, March 15, 2025). This means that one area of interest could be a better system of outreach or public awareness that connects the goals directly to the people meant to enact them. It has also been scientifically proven that improved education about climate change leads to increased support for climate-friendly policies (Dabla-Norris & Khalid, 2023).

Beyond the limitations expressed in the methodology, some also exist in the implications derived from this research. The most important factor to note is the forced simplification of regression models due to a few reasons. The biggest reason why much of the detail was abstracted was due to the lack of a sample size. Since many of the policies that drive impacts such as EV and solar panel incentives have only come into effect around 2019, it becomes difficult to understand whether trends will continue linearly as has been generally assumed. The methodology itself is also hindered in achieving complete efficiency because much of the data was missing key values and had to be cleaned out of the dataset entirely. Since the dataset also draws from multiple sources to reduce bias, some reference calculations such as those regarding EV percentage of cars on the road yield different values based on which source of data is used. In the future, more data will be available and trends may be more accurate compared to what is available as of April 2025. Future studies should also include direct collection of data rather than relying on what is publicly available as well as multivariable regression if more types of data can be maintained which would likely be more accurate of an indicator for a time-series prediction. In addition to having more data, a better statistical representation of demographic data in the regression models could serve to further highlight potential solutions.

Works Cited

Alvarez-Palau, E. J. “Transport and Urban Growth in the First Industrial Revolution.” Yale Economics, 30 Sept. 2020, https://economics.yale.edu/sites/default/files/marketaccessgrowth_presteam_sept272020.pdf.

Barr, H., et al. RESPONSE ACT. New Jersey Department of Environmental Protection, 15 Oct. 2020, https://dep.nj.gov/wp-content/uploads/climatechange/nj-gwra-80x50-report-2020.pdf.

Barrage, L., and W. D. Nordhaus. “Policies, Projections, and the Social Cost of Carbon: Results from the DICE-2023 Model.” NBER, 10 Apr. 2023, https://www.nber.org/papers/w31112.

Bonannella, C., et al. “Biomes of the World under Climate Change Scenarios: Increasing Aridity and Higher Temperatures Lead to Significant Shifts in Natural Vegetation.” PeerJ, 2023, https://doi.org/10.7717/peerj.15593.

Cutler, D., and G. Miller. “The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States.” Demography, vol. 42, no. 2, 2005.

Dabla-Norris, E., and S. Khalid. “Public Support for Climate Change Mitigation Policies: A Cross-Country Survey.” International Monetary Fund, vol. 2023, no. 223, 2023, https://www.elibrary.imf.org/view/journals/001/2023/223/article-A001-en.xml.

Graven, H., R. Keeling, and J. Rogelj. “Changes to Carbon Isotopes in Atmospheric CO_2 over the Industrial Era and into the Future.” Global Biogeochemical Cycles, vol. 34, 2020, https://doi.org/10.1029/2019GB006170.

Nguyen, T., et al. “Combining Datasets to Improve Model Fitting.” arXiv, May 2023, https://arxiv.org/pdf/2210.05165.

Obschonka, M. “Research: The Industrial Revolution Left Psychological Scars That Can Still Be Seen Today.” Harvard Business Review, Mar. 2018, https://hbr.org/2018/03/research-the-industrial-revolution-left-psychological-scars-that-can-still-be-seen-today.

OER Commons. “Statewide Dual Credit World History, The Period of Revolution 1650–1871 CE, Chapter 8: Industrial Revolution.” 2018, https://oercommons.org/courseware/lesson/87910/student/?section=1.

Ritchie, H., and M. Roser. “CO_2 Emissions.” Our World in Data, 2020, https://ourworldindata.org/co2-emissions.

Rocque, R. J., et al. “Health Effects of Climate Change: An Overview of Systematic Reviews.” BMJ Open, 2021, https://doi.org/10.1136/bmjopen-2020-046333.

Ruppar, T. “Meta-Analysis: How to Quantify and Explain Heterogeneity?” European Journal of Cardiovascular Nursing, vol. 19, no. 7, Oct. 2020, pp. 646–652, https://doi.org/10.1177/1474515120944014.

Singh, B. P., et al. “Efficient Dataframe Systems: Lazy Fat Pandas on a Diet.” arXiv, 2024, https://arxiv.org/html/2501.08207v1.

The Conference Board. “Report: Solar Power Is Cheaper than Ever—But Overreliance Risks Destabilizing US Electricity Markets and Undermining Energy Investments.” PR Newswire, 21 Feb. 2024, https://www.prnewswire.com/news-releases/report-solar-power-is-cheaper-than-everbut-overreliance-risks-destabilizing-us-electricity-markets-and-undermining-energy-investments-302067341.html.

U.S. Energy Information Administration (USEIA). U.S. Energy Information Administration - EIA - Independent Statistics and Analysis. https://www.eia.gov/state/?sid=NJ.

United States Environmental Protection Agency. “Causes of Climate Change.” EPA, 25 Mar. 2025, https://www.epa.gov/climatechange-science/causes-climate-change.

van Neuss, L. “Why Did the Industrial Revolution Start in Britain?” University of Liège, 2015, https://orbi.uliege.be/bitstream/2268/188592/1/Why

Characterizing Discussion Topics on CCUS in Twitter Data via Textual and Sentiment Features By Emily Ma

Abstract

This paper plans to analyze tweets on carbon capture, utilization and storage (CCUS) to explore the discussion topics among the tweets. After processing and cleaning a large data set of tweets, Latent Dirichlet Allocation(LDA) was used to identify dominant topics and their distributions. Visualizing the results with t-distributed Stochastic Neighbor Embedding (t-SNE) showed clear topic clusters with some overlapping discussion topics. The results provide insights into public opinion on CCUS, offering stakeholders and policy makers a better understanding about the current discussion topics on CCUS.

1 Introduction

Carbon capture, utilization and storage (CCUS) captures carbon dioxide (CO2) emissions from industrial sources and reuses or stores them underground. It has gained attention as a promising tool for climate change as it offers “low-carbon, zero-carbon, and even negative-carbon solutions” (Li). Public opinion is important for policy makers to understand before they make decisions surrounding tech implementations or policy changes. Social media platforms like Twitter provide a solid view of the opinions and discussions surrounding CCUS. Manually annotating large amounts of data is extremely time consuming, which is why this project aims to increase efficiency with the help of topic modeling. By applying a topic model on tweets, this project aims to discover whether or not there are patterns in discussions about CCUS and where the majority lies in these discussions. Identifying these details could allow stakeholders to understand how CCUS is perceived and change their strategy based on the results.

2 Literature Review

Through recent advancements in natural language processing, it is now possible to analyze public opinions and their trends using social media data, survey data, and much more. Topic modeling has surfaced as a widely used tool to uncover themes in large text data sets. In a paper published in 2023 by Lerksuthirat et al., sentiment analysis and topic modeling was applied on a data set containing tweets expressing opinions on cannabis-containing foods after legalization in Taiwan(Lerksuthirat et al.). This study revealed that the public expressed negative sentiments after the policy and identified distinct groups of concerns regarding health and regulatory issues. In another study conducted by Huang et al. in 2023, LDA topic modeling was used on Derwent patent data to outline the evolution of CCUS technologies(Huang et al.). They illustrated the path that technologies surrounding CCUS took over time with the help of topic modeling. While Huang et al. used topic modeling to examine patent data as a way to understand the growth pattern of CCUS, Lerksuthirat et al. used topic modeling to analyze public sentiment and discussion topics surrounding a policy change. Despite the large amount of research

involving topic modeling and CCUS, there is a lack of research on applying topic modeling to current opinions about CCUS within social media platforms. Building on the methodologies of these two studies, this paper aims to discover topics of discussion on CCUS using Twitter data.

3 Data & Methodology

3.1 Data

The data set used contains tweets in any language supported by Twitter in the months January to March 2023 with any mention of CCUS. The scraping process was done with the official Twitter API in Python. After being machine translated into English, the tweets were manually annotated. The initial tweets contained special characters, URLs, and mentions. The data set was found to be sufficient for the model due to its balanced amount of positive and negative tweets, as shown in Figure 1 illustrating the sentiment distribution of the tweets.

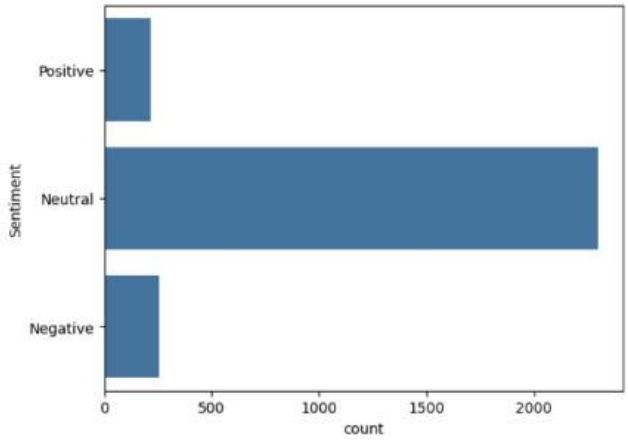


Figure 1: Bar graph of amount of positive, negative, and neutral tweets.

3.2 Processing Data

In order to prepare the tweets for LDA, the tweets were cleaned using regular expressions in Python to remove noise and ensure consistency. The tweets were cleaned from whitespace, URLs, mentions, apostrophes, and all non-alphabetic characters. After the tweets were cleaned, a bi-gram natural language tokenizer was used on them to prepare for the transformation into a bag-of-words document, a list of words and their frequencies, which is the type of data that LDA works with.

3.3 Methodology

LDA was chosen among other topic modeling algorithms because of its ability to assign each tweet with multiple topics, similar to real life scenarios. Using the processed data, a dictionary was created to map each word to an integer because LDA works with integers, not raw text. The dictionary was then used to create a corpus, which shows which words appear in the tweets and how often they appear. After processing the tweets, topic modeling was

implemented using LDA, following publicly available tutorials(GeeksforGeeks). The performance of the model will be determined with coherence scores. The c_v coherence score ranges from 0 to 1, with higher scores representing greater model coherence, and lower scores indicating that there is more noise in the topics generated. To visualize the data, we will use t-SNE, a dimensionality reduction method. This is because of t-SNE's ability to simplify the high-dimensional topic distributions of each tweet into 2D so that they can be plotted as a scatter plot(Wikipedia contributors).

3.3 Initial Results

After implementing LDA with 10 topics, a coherence score of ~0.39 was outputted. Looking at the words sorted into each topic, a majority of the words in each sorted topic were more general words such as 'they', 'be', and 'say', indicating that topics generated may be inaccurate and affected by noise. In addition, upon visualizing the distribution of the topics using a scatter plot created with the help of t-SNE, topics 6 and 7 dominated the graph while the other topics mixed with these two topics as shown with Figure 2.

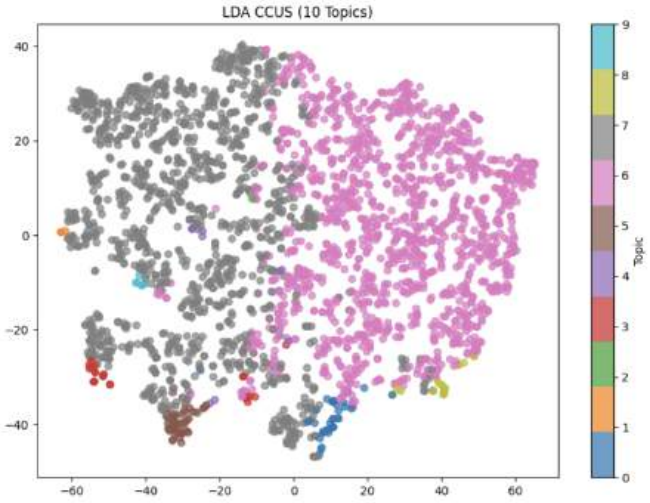


Figure 2: Clusters of documents based on the LDA model with 10 topics.

3.3 Improvement Methods

The convoluted results of the above LDA model may be due to the amount of topics, a key hyperparameter of LDA. As noted in a journal published by Gan Jingxian and Yong Qi regarding the selection of optimal topics for LDA models, LDA will produce topics that are too general when forced to generate less topics than the optimal number(Gan and Qi). Likewise, when forced to generate more topics than the optimal amount, it creates similar topics or useless topics. Therefore, the topic count for the LDA models were set to 5, 4, 6, and 15 topics receptively and compared them based on their coherence scores. As shown in Figure 3, the LDA model with 5 topics was the most accurate with a coherence score of about 0.41.

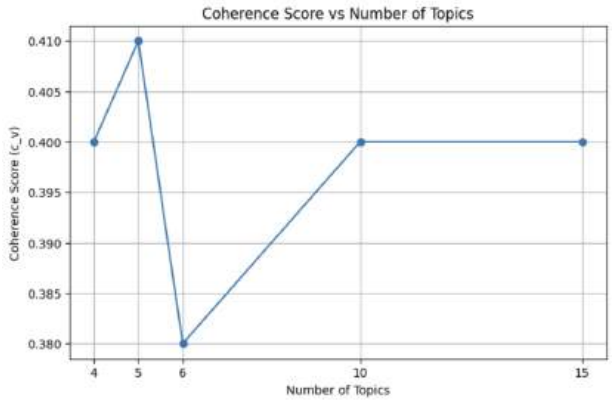


Figure 3: Plot of coherence scores

4 Results

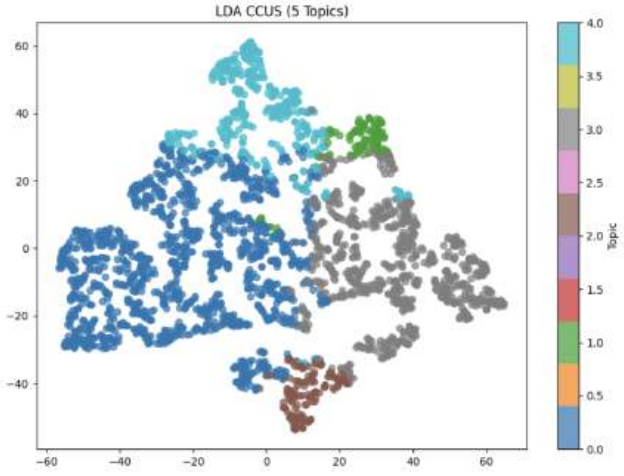


Figure 4: Clusters of documents based on the LDA model with 5 topics.

The five topics generated by the model, identified with the help of ChatGPT("ChatGPT - Remove Numerical Values"), are color-coded in Figure 4 as:

- **Topic 0**: General talk about CCUS (Dark Blue)
- **Topic 1**: CCUS and other clean technologies (Green)
- **Topic 2**: CCUS as a future climate solution (Brown)
- **Topic 3**: Technological viewpoint on CCUS (Gray)
- **Topic 4**: Personal opinions on CCUS (Light Blue)

The distribution of the tweets was then visualized in a scatter plot with t-SNE in Figure 4. The model displays a clear separation of topics with some of the tweets bordering two to three topics, which demonstrates the model's ability to identify topics within the tweets. The dominant topic of each tweet was determined and used to form Figure 5, a bar graph representing the distribution of tweets that identified as positive, negative, and neutral according to the prelabeled sentiment column. The large amount of neutral tweets were expected as the original data was

also largely labeled as neutral. Looking at the amount of negative tweets compared to the positive tweets, it can be said that much of the talk about CCUS in the future, CCUS and other clean technology, and technological components of CCUS is positive while general talk and personal opinions about CCUS are more negative.

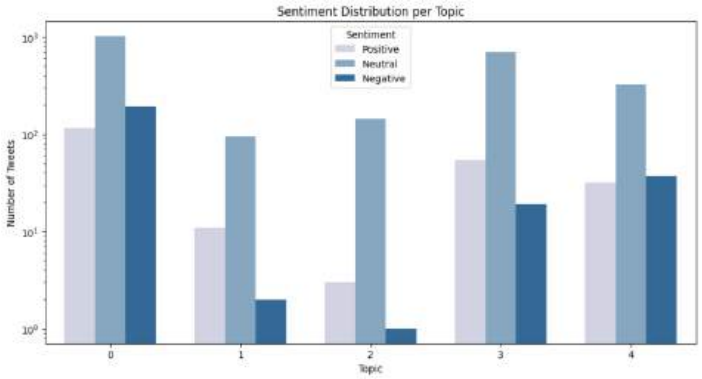


Figure 5: Bar graph of sentiment distributions

5 Conclusions

This paper identified five main discussion topics surrounding CCUS. The optimal number of topics for the LDA model was determined with the use of coherence scores, leading to a more accurate model. From the sentiment distributions per topic, the public demonstrates a more negative opinion when discussing CCUS in general while discussions around incorporating CCUS with other clean tech and the implications of CCUS in the future demonstrated a more positive sentiment. The results show that topic modeling could be a potential way to quickly determine opinions surrounding CCUS and could be potentially beneficial as a tool for people who base their actions on public opinion. In the future, the dataset used could be improved with a survey asking questions about concerns, benefits, and opinion as a majority of tweets on CCUS are neutral and the data would be more precise and contain more insights with a survey.

Works Cited

Li, Hu. “Advancing ‘Carbon Peak’ and ‘Carbon Neutrality’ in China: A Comprehensive Review of Current Global Research on Carbon Capture, Utilization, and Storage Technology and Its Implications.” ACS Omega, vol. 8, no. 45, Nov. 2023, pp. 42086–101. https://doi.org/10.1021/acsomega.3c06422.

Lerksuthirat, Tassanee, et al. “Sentiment and Topic Modeling Analysis on Twitter Reveals Concerns Over Cannabis-Containing Food After Cannabis Legalization in Thailand.” Healthcare Informatics Research, vol. 29, no. 3, July 2023, pp. 269–79. https://doi.org/10.4258/hir.2023.29.3.269.

Huang, Liangchao, et al. “Evolution of CCUS Technologies Using LDA Topic Model and Derwent Patent Data.” Energies, vol. 16, no. 6, Mar. 2023, p. 2556. https://doi.org/10.3390/en16062556.

Wikipedia contributors. T-distributed Stochastic Neighbor Embedding. 24 May 2025, en.wikipedia.org/wiki/T-distributed_stochastic_neighbor_embedding.

GeeksforGeeks. “Topic Modeling Using Latent Dirichlet Allocation (LDA).” GeeksforGeeks, 23 July 2025, www.geeksforgeeks.org/nlp/topic-modeling-using-latent-dirichlet-allocation-lda.

Gan, Jingxian, and Yong Qi. “Selection of the Optimal Number of Topics for LDA Topic Model—Taking Patent Policy Analysis as an Example.” Entropy, vol. 23, no. 10, Oct. 2021, p. 1301. https://doi.org/10.3390/e23101301.

“ChatGPT - Remove Numerical Values.” ChatGPT, chatgpt.com/share/68827f31-5dd4-8001-bdae-df89b3289b66.

AI-Powered Multilingual STEM Learning: A Feasibility Study By Siri Mutalik Desai

Abstract

While over 7,000 languages are spoken across the globe today, English remains the primary language of science, technology, engineering, and mathematics (STEM). Although the goal of this linguistic uniformity was to streamline scientific communication, it has also created significant barriers. Research and policy highlight the widespread exclusion of non-English-speaking students and researchers from STEM pathways due to limited multilingual instruction, which results in underrepresentation and lower academic achievement in linguistically marginalized communities. This paper reviews the impacts of English dominance on scientific research and learning, outlines the barriers in multilingual STEM education, and discusses the potential of artificial intelligence (AI) to improve multilingual STEM instruction. A case study of Polyglot STEM Buddy, an open-source AI chat-based application developed by the author to teach STEM in multiple languages, is presented as a model for inclusive STEM education.

Introduction

More than 7,000 languages are spoken worldwide, yet since 1967, English has dominated global scientific communication (Henry et al.). While this common language has streamlined communication, it has also limited access to STEM for non-English-speaking communities. Major English-language journals often overlook scientific research published in other languages, which limits the visibility of important findings. This lack of linguistic diversity in STEM fields has stifled innovation and exacerbated global inequality.

This paper addresses the question: How can AI play a role in advancing multilingual STEM education for non-English-speaking students? Here, "non-English-speaking" refers to students or communities whose primary home language is not English. This definition encompasses multilingual learners, who are "students developing proficiency in multiple languages" (WIDA). This paper argues that AI-based multilingual tools are essential to help dismantle language barriers and promote equity in STEM education. It reviews literature on English dominance, highlights the need for multilingual STEM education, and explores how AI can power multilingual educational tools. The paper concludes by introducing Polyglot STEM Buddy, an open-source AI chatbot designed by the author to teach STEM to students in their home languages. Note that "home language" refers to the language most commonly spoken at home by a speaker's family.

Literature Review

The Consequences of English Dominance in STEM

Many non-English-speaking scientists are marginalized, particularly in publishing, due to the dominance of the English language. The *Journal of Science Policy and Governance* reports

that scholarly work published in non-English languages is frequently excluded from high-impact citations, which may negatively impact research visibility and career advancement for these researchers (Henry et al.).

One clear example of this exclusion occurred during the COVID-19 pandemic, when English-speaking researchers overlooked important studies originally published in Chinese, potentially delaying the communication of important knowledge during a global health crisis (Xiang). Such incidents highlight how English dominance not only damages researchers' careers but also hinders the collaboration and scientific progress necessary for societal growth.

Limitations of Relying on Translation

One proposed solution to the lack of multilingualism in STEM is the use of translation services, which translate academic papers from one language to another. However, this solution faces two significant challenges: (a) a lack of highly technical translators, and (b) financial barriers. First, there exists a great shortage of highly technical translators who have both complex knowledge of STEM and the cultural fluency necessary for accurate translation. Second, the cost of translation services can be a significant barrier, with fees reaching as high as $10,000 per manuscript (3,500–10,000 words). Furthermore, the lack of transparency in pricing practices puts non-native English authors at a disadvantage, requiring negotiations in a language they are less comfortable with (Henry et al.). This time-consuming, expensive process creates barriers in the translation, communication, and distribution of important scholarly work.

Consequently, human-provided translation services are not a simple, effective solution to the lack of multilingualism in STEM. Rather, this issue must be resolved by considering its underlying factors. One such factor is the lack of multilingualism within STEM education, which not only limits access to STEM for non-English-speaking communities but may also contribute to the shortage of qualified technical translators.

The Necessity and Benefits of Multilingual STEM Education

Multilingual STEM education offers numerous benefits, including improved cognitive development and academic achievement. Students learning STEM in their home languages demonstrate better comprehension and retention, which likely arises from increased linguistic comfort (Mitchell; UNESCO). These findings are supported by Cummins' Interdependence Hypothesis, which proposes that skills and knowledge acquired in one language will contribute to learning and proficiency in another through a common underlying proficiency framework (Goodrich and Lonigan). This theory explains why improving multilingual STEM education can improve overall academic outcomes and second language acquisition. The case for multilingual STEM education is further strengthened by cognitive load theory, which suggests that reducing linguistic demand frees mental resources for complex ideas (Chen and Martinez). In other words, when linguistic demand is reduced (by providing instruction in a familiar language), students have more mental resources available for and are able to more easily learn complex ideas (like those seen in STEM).

However, students from migrant and multilingual backgrounds generally underperform in STEM, mostly due to language barriers and a scarcity of multilingual instructors (Mitchell; Smit et al.). These language barriers are significant: over 40% of students worldwide lack access to education in a language they fully understand; in some countries, this figure rises to 90% (UNESCO). By contributing to the underperformance of multilingual learners in STEM, these disparities have the potential to contribute to the underrepresentation of these students in STEM careers, which has long-term consequences for workforce diversity (McInerney and Mills).

Artificial Intelligence and Educational Technology

Artificial intelligence (AI) and its integration into education via the educational technology (EdTech) sector offer a unique method to advance multilingual STEM education. AI-powered platforms can deliver standards-aligned instruction across multiple languages, personalizing lessons to a learner's proficiency and age (MRCC Technology). AI-powered features such as instant assessments, culturally and technically accurate STEM translation, learner dashboards, and voice-based learning can help strengthen both language skills and STEM understanding. Since the exposure of STEM to multilingual learners has been linked to improved academic outcomes (Mitchell), AI's ability to teach STEM to multilingual learners may help improve academic outcomes.

In comparison to general translation tools—which educators may use to translate learning materials for multilingual learners—educational AI stands out for many reasons. First, general translation tools, such as Google Translate, lack the contextual and cultural awareness that is often necessary for complex translations. By contrast, educational AI tools like large language model-powered translators possess contextual and cultural awareness, which allows for more natural, human-sounding, and accurate translations of learning material ("Best AI Translators for Education"). By maintaining dialogue and adapting lessons based on learner response, EdTech AI-based tools have the potential to help improve comprehension and understanding of STEM concepts.

Counterarguments and Limitations of AI in Education

While AI-powered educational tools offer significant benefits, several limitations must be acknowledged. First, AI models may perpetuate biases present in training data, which may reinforce stereotypes or cultural misconceptions in STEM explanations ("When AI Gets It Wrong"). Second, AI lacks the "vital human qualities crucial for effective learning experiences" (Giray)—in other words, AI cannot connect with students on an emotional, human-to-human level. Consequently, AI tools may be limited in their ability to personalize lessons because they lack the emotional intelligence of human teachers. Finally, over-reliance on AI tutoring may harm users by lowering cognitive activity, which may negatively impact STEM skill development (Kosmyna et al.).

These limitations highlight the importance of positioning AI tools as supplements to, rather than replacements for, human instruction. Despite these limitations, the potential benefits

of strategically and carefully implemented AI multilingual education tools outweigh the risks and ought to be considered.

Methodology

This pilot study represents exploratory research designed to assess the feasibility of an AI-based educational tool (Polyglot STEM Buddy) developed by the author. The study focuses on using formative evaluation techniques to analyze user reception and perceived usability. The study's purpose is to determine whether Polyglot STEM Buddy warrants further development and formal testing.

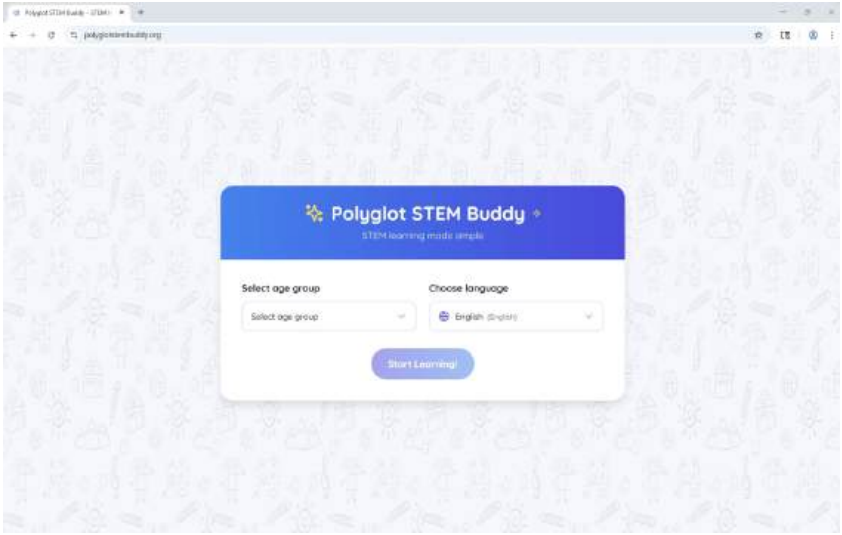
Application Design

Polyglot STEM Buddy, a web-based application, was conceptualized and developed by the author. Translation services were implemented via OpenAI's GPT-4o mini model via Microsoft Azure's OpenAI API, which mitigated the need for manual translation validation and ensured consistent, professional-quality translations across all supported languages. OpenAI's GPT-4o mini model is the more cost-efficient, lightweight variant of their GPT-4o model. The decision to utilize the existing API allowed the research to focus on the usability of the application rather than translation mechanics.

The author's role as a developer and a researcher enabled the author to closely monitor user responses and make timely improvements based on user feedback. User feedback was anonymized to ensure objectivity.

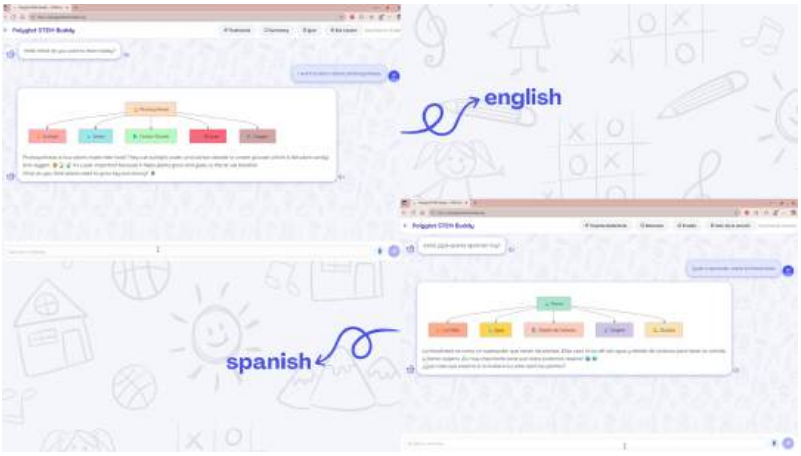
Figure 1 depicts the application's landing page, where age group and language can be selected; and Figure 2 displays a side-by-side comparison of Polyglot STEM Buddy in a conversation about photosynthesis in English (left) and Spanish (right).

Figure 1: Polyglot STEM Buddy Landing Page



:

Figure 2: Polyglot STEM Buddy Chatbot



The features in Figures 1 and 2 depict the main features tested during the pilot study: the ability to select age/language, and the chatbot's ability to engage users. Non-depicted features that were tested include flashcard generation, quiz generation, and summary generation.

Research Design and Participant Sampling

The Polyglot STEM Buddy application was evaluated through a mixed-methods research design. Fifteen participants were recruited using purposive sampling to ensure a diverse array of perspectives, consisting of high school students, parents of young children, STEM educators, and multilingual learners. No sensitive or identifying data were collected, and all participants provided informed consent or assent prior to participation.

Participants completed a survey after interacting with Polyglot STEM Buddy. Survey questions focused on evaluating the app's functionality, usability, and educational value.

Data Collection and Evaluation Metrics

Data collection combined qualitative and quantitative methods. Qualitative insights came from open-ended surveys accessible via the app's "Feedback" button and social media outreach on Instagram and WhatsApp. Multilingual STEM learners self-assessed content accuracy. User satisfaction was measured with Likert-scale questions focused on usability, clarity of explanations, and UX/UI design. Thematic and content analyses of the collected data informed iterative improvements and provided insights into the application's educational potential.

Ethical Disclosure

Polyglot STEM Buddy was both designed and tested by the author. User data was collected and analyzed with transparency and care to accurately represent participant experiences. The anonymization and independent verification of user feedback were utilized to mitigate potential conflicts of interest. The survey questions were worded neutrally in order to avoid skewing data through "leading questions."

Results and User Feedback

Feedback was gathered from a pilot study with 15 users—including students, parents, and educators—after interacting with Polyglot STEM Buddy, a web-based application. It is important to note that the results of the survey only reflect user satisfaction and experience, not educational efficacy or learning outcomes.

Responses were summarized quantitatively and qualitatively in Table 1. Overall helpfulness received strong positive ratings (93%), as did App Navigation and UX Design, which scored 100% positive. Users described the app as engaging with a user-friendly interface. No translation errors were reported, and testing spanned all age groups and supported languages.

The efficacy of follow-up questions earned slightly lower positive feedback (80%), with users requesting more specific and engaging prompts. Some participants noted issues with audio and visual rendering (such as incorrect fraction displays), leading to a lower satisfaction rate (60%) in this area. Despite this, the majority (87%) said they would recommend the app to others, including younger siblings and peers.

These responses suggest that the application's core design features are well-received by a diverse user base, which indicates potential for broader interest and further study. The next research phase will incorporate more formal studies, involving control/experimental groups, standardized learning assessments, and a longitudinal aspect in order to accurately measure educational efficacy.

Table 1: Summary of User Feedback on Polyglot STEM Buddy

Aspect Evaluated	Positive Responses (%)	Neutral Responses (%)	Negative Responses (%)	Key Qualitative Comments
App Navigation/UX Design	100%	0%	0%	"Accessible and easy to use"; "Cool interface!"
Explanation Clarity	93%	7%	0%	"Shorten summary"; "Overlap in STEM concepts"
Follow-Up Question Efficacy	80%	20%	0%	"Add clickable individual elements"; "More specific follow-up questions"
Help Delivered	93%	7%	0%	"Helpful!"; "Add option to be in-depth or not"
Audio/Visual Aid Satisfaction	60%	40%	0%	"Fractions rendered wrong"
Likelihood to Recommend the App	87%	13%	0%	"Would recommend to younger siblings"; "Useful to students"

Note: Feedback was collected on a 1-5 scale, with 1-2 considered negative, 3 neutral, and 4-5 positive. Percentages reflect the proportion of respondents expressing positive, neutral, or negative views for each category. UX refers to user experience.

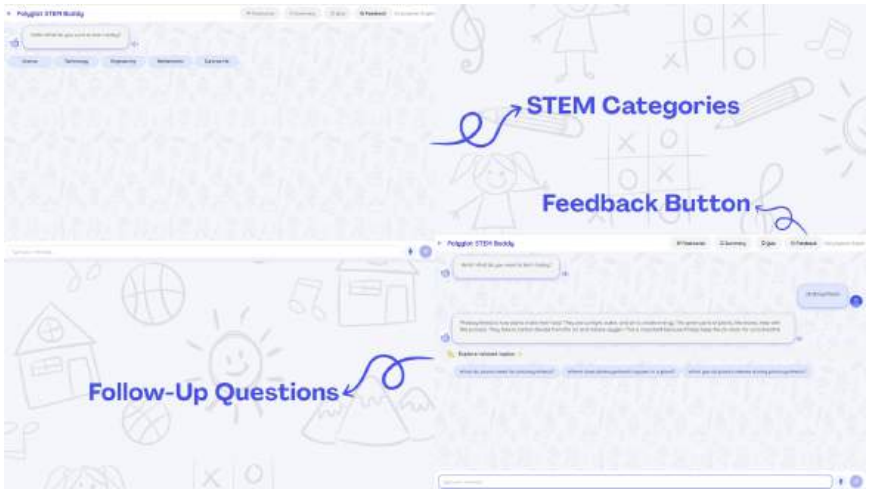
Tool Enhancements Based on Feedback

The author's unique perspective as both developer and researcher enabled rapid, informed improvements during the study by directly applying firsthand user insights. Based on user feedback, several improvements were implemented to improve user experience and app functionality:

- Introduction of five clickable topic buttons (STEM category buttons and a "Surprise Me" option) at the start of a conversation to help ease the user into a lesson.
- Addition of follow-up question buttons to maintain engagement throughout the lesson.
- Replacement of an "Exit Lesson" button with a "Feedback" button, providing an easy avenue to report bugs and promote user-driven improvements.
- Resolution of audio playback issues to improve accessibility for users relying on audio/voice mode.

In Figure 3, the category buttons, follow-up question buttons, and the "Feedback" button features are visible.

Figure 3: Polyglot STEM Buddy Added Features



These features are expected to improve user satisfaction, though further testing would be necessary to verify the claim. All features can be tested at https://polyglotstembuddy.org/.

Discussion

Translation is often seen as a simple fix for language barriers in global scientific communication. However, financial hurdles and a shortage of qualified, culturally aware, highly technical translators impede effective implementation. Clearly, an effort to advance multilingualism in STEM is necessary—and this effort must begin with education.

AI tools like Polyglot STEM Buddy offer methods to advance multilingual education by personalizing lessons according to learners' age and language proficiency. These AI tools, integrated appropriately with human instruction, have the potential to (a) make STEM education

more accessible for non-English-speaking communities/learners, and (b) to help improve academic outcomes for multilingual learners. In doing so, AI may help increase representation and diversity in the STEM workforce.

These findings extend beyond individual outcomes. Diversity has long been linked to higher innovation rates and unique problem-solving approaches (Crawford). By looking at Cummins' Interdependence Hypothesis and the cognitive load theory, and by working with students, parents, and educators, multilingual STEM education tools can be developed to potentially help improve academic outcomes, diversify the STEM workforce, and thus help increase innovation.

Critical Evaluation of the Current Application

Although user feedback was positive, there are several limitations to Polyglot STEM Buddy. First, the application's reliance on a general AI model (GPT-4o mini) rather than a model trained on standard-aligned content may limit its educational value, as its current knowledge is comparable to that of other general LLMs and may contain inaccuracies. Moreover, due to a lack of standardization of the content, it may be difficult to integrate Polyglot STEM Buddy into a classroom, as it may not align with formal education standards. Additionally, a lack of features such as progress tracking or user creation may make it less useful for educators.

Ethical Considerations

This study addresses four primary ethical concerns: data privacy, AI content bias, digital equity, and cultural sensitivity. Regarding data privacy, the application complies with current data protection guidelines through anonymized user feedback and secure AI hosting. To strengthen data privacy in the future, later versions of Polyglot STEM Buddy must incorporate stronger safeguards in order to protect vulnerable populations, such as young children.

Another concern that the study addresses is the risk of AI bias. Although mitigated by the use of the trusted GPT-4o mini model and the use of clear instructions meant to keep content appropriate, the challenge of AI bias is still present. AI models like GPT-4o mini, trained on primarily English-language datasets, may inadvertently privilege Western ideologies or inaccurately represent foreign cultural contexts. In order to identify and correct these biases, regular collaboration with native speakers is necessary.

Digital equity is another ethical concern—this concern extends past bandwidth limitations to encompass access to devices and technology. Although Polyglot STEM Buddy is planned to include offline functionality and low-bandwidth optimization to mitigate accessibility barriers, learners located in certain regions may lack access to the infrastructure necessary to utilize the application. The issue of digital equity must be addressed through policy interventions and NGO work to increase access to technological infrastructure. Until such access is improved, the risk persists that AI-powered educational tools may inadvertently widen existing educational gaps (like the one caused by the digital divide) rather than close them.

Finally, the risk of cultural insensitivity must be addressed in order to ensure translations remain technically and culturally accurate, and to help create a safe, welcoming learning environment for learners. Ongoing collaboration with multilingual communities is necessary to maintain a respectful atmosphere and to contextualize STEM concepts in a cultural manner.

Together, these considerations highlight the responsibility that accompanies the use of AI in EdTech contexts.

Comparative Analysis

To assess the benefits Polyglot STEM Buddy offers relative to its competitors, a comparative analysis chart was created. This comparison identifies strengths and weaknesses in similar EdTech tools, highlighting areas where Polyglot STEM Buddy does well and where it can improve. By examining features such as language support, AI capabilities, user experience, and transparency, this analysis clarifies the application's current state and guides future development to better meet the needs of multilingual STEM learners.

Table 2: Polyglot STEM Buddy vs. Existing Solutions

Feature	Polyglot STEM Buddy	Khanmigo	ChatGPT Study Mode	YouLearn.ai	Comments
Multilingual STEM Support	5 languages, with additional languages in development	11 languages; focus on English, Spanish, Hindi, and Portuguese	59 languages	21 languages	ChatGPT currently leads; Polyglot STEM Buddy aims to expand multilingual support to remain competitive
AI Interactivity	Powered by GPT-4o mini; asks follow-up prompts; conversational tone	GPT-powered conversations, with structured guidance	Conversation Q&A, with file/image upload support	Tutor chat; explanations about user-uploaded content	Khanmigo excels in learner guidance; Polyglot STEM Buddy will grow plans to develop similar features
UX Features	Topic buttons, voice/audio feature, follow-up questions, STEM visuals, flashcards, summary, quizzes; tailored for younger learners	Tailored for educators and learners; includes teacher dashboards and assignments	Chat-based, file/image upload, voice/audio	Summary, flashcards, quizzes, voice mode; tailored toward older learners (high school, college)	All applications enable active learner participation and iterative design

Content Quality	Generic model knowledge	Khan Academy curricula-aligned, educator-vetted	Generic model knowledge	Variable, based on user/PDF/video upload; generic model knowledge	Khanmigo currently leads in alignment with curricula; in the future, Polyglot STEM Buddy plans to collaborate with educators to enhance content quality
Open-Source Availability	Yes	No	No	No	Polyglot STEM Buddy is the only open-source tool, enhancing transparency and improving accessibility

Note: This table compares Polyglot STEM Buddy with other educational AI platforms (Khanmigo, ChatGPT's Study Mode, and YouLearn.ai) based on language support, AI capabilities, user interaction features, content quality, and transparency. This analysis is based on publicly available feature descriptions as of August 2025 and direct user testing for Polyglot STEM Buddy. Differences in platform design goals and user bases should be considered when interpreting results.

Limitations and Future Work

While promising, this study faces several limitations. The small, non-random sample ($n = 15$) restricts the generalizability of findings. In order to thoroughly evaluate the application's impact, future formal testing in classroom environments is necessary. Furthermore, due to the lack of statistical significance testing, reported differences in user satisfaction may not represent meaningful distinctions between features, which limits the reliability of comparative evaluations.

Ongoing ethical considerations—such as protecting user privacy, monitoring potential AI biases, digital equity, and cultural sensitivity—must be continually addressed in future work through collaboration with multilingual students, parents, educators, and communities.

Beyond technical challenges, several non-technical barriers may hinder the wider adoption of multilingual AI education tools (Ansari et al.). Educational systems often require extensive pilot testing before implementing new technologies, which means institutional approval processes can take months. Limited technological infrastructure and varying levels of digital literacy may obstruct the ability of systems to adopt new technologies (OECD; Shoemaker DeMio). Additionally, in order to successfully use these tools, teachers must be trained in integrating the technology into the classroom in a manner that mitigates the risk of over-reliance on AI (Herold). Finally, tools must be properly vetted and comply with strict privacy protection regulations in order to protect student privacy (U.S. Department of Education).

Future research will prioritize rigorous experimental design (with proper sampling techniques, control/experimental groups, and standardized learning assessments), longitudinal studies (to track learning outcomes over longer periods of time), statistical analysis (to determine the significance of findings), and collaboration with educators (to align content to educational standards). Additionally, future work will prioritize expanding language options based on user demographics and fostering closer collaboration with educators, students, and parents. Development goals include improving offline accessibility, adding gamification features to enhance learner motivation, and making Polyglot STEM Buddy AI-agnostic to increase accessibility for developers and broaden deployment opportunities.

Conclusion

The dominance of English in STEM research hinders equity, accessibility, and innovation. Multilingual STEM education is necessary to improve multilingualism in the global scientific community—and AI tools in EdTech have the power to transform multilingual STEM education. Tools like Polyglot STEM Buddy, created to teach STEM to children in the language they are most comfortable with, highlight the crucial role that AI can play in helping to advance multilingual STEM education for non-English-speaking students.

This pilot study establishes a foundation for future rigorous research and provides a replicable framework for AI-powered EdTech tools that prioritize language, culture, and precision. Future studies must employ controlled experimental designs with proper sample sizes, standardized learning assessments, and longitudinal aspects to determine the impact of such AI tools on academic achievement. The finding that 87% of users would recommend the application, despite technical limitations, suggests strong demand for accessible multilingual STEM education tools, indicating both market viability and educational necessity.

Acknowledgements

This work was independently undertaken, with informal user feedback from educators and peers.

Works Cited

Ansari, Marium, et al. “Barriers to Educational Technology Adoption: Navigating Challenges in Integration.” *Qlantic Journal of Social Sciences*, vol. 5, no. 3, 30 Sept. 2024, pp. 240-47, https://doi.org/10.55737/qjss.123732538. Accessed 9 Aug. 2025.

“Best AI Translators for Education Built for Classrooms.” *MachineTranslation.com*, 21 May 2025, www.machinetranslation.com/blog/best-ai-translators-for-education. Accessed 7 Aug. 2025.

Crawford, Krysten. “Stanford study provides evidence of the 'diversity-innovation paradox' in academia.” *Stanford Graduate School of Education*, 24 Apr. 2020, ed.stanford.edu/news/stanford-study-provides-evidence-diversity-innovation-paradox-academia. Accessed 7 Aug. 2025.

Dietrich, Sandy, and Erik Hernandez. “What Languages Do We Speak in the United States?” *United States Census Bureau*, 6 Dec. 2022, www.census.gov/library/stories/2022/12/languages-we-speak-in-united-states.html. Accessed 23 July 2025.

“Digital divide in education.” *OECD*, www.oecd.org/en/topics/sub-issues/digital-divide-in-education.html. Accessed 9 Aug. 2025.

Giray, Louie. “Ten Myths About Artificial Intelligence in Education.” *Higher Learning Research Communications*, vol. 14, no. 2, 2024, pp. 1-12. *ERIC*, https://doi.org/10.18870/hlrc.v14i2.1508. Accessed 7 Aug. 2025.

Goodrich, J. Marc, and Christopher J. Lonigan. “Language-independent and Language-specific Aspects of Early Literacy: An Evaluation of the Common Underlying Proficiency Model.” *Journal of Educational Psychology*, vol. 109, no. 6, Aug. 2017, pp. 782-93. *PubMed*, https://doi.org/10.1037/edu0000179. Accessed 9 Aug. 2025.

Henry, Kaylee, et al. “A Call to Diversify the Lingua Franca of Academic STEM Communities.” *Journal of Science Policy & Governance*, vol. 18, no. 03, 30 Aug. 2021. *ERIC*, https://doi.org/10.38126/jspg180303. Accessed 9 Aug. 2025.

Herold, Benjamin. “Poor Students Face Digital Divide in How Teachers Learn to Use Tech.” *Education Week*, Editorial Projects in Education, 12 June 2017, https://www.edweek.org/leadership/poor-students-face-digital-divide-in-how-teachers-learn-to-use-tech/2017/06. Accessed 9 Aug. 2025.

“How to Change Your Language Setting in ChatGPT.” *OpenAI*, July 2025, help.openai.com/en/articles/8357869-how-to-change-your-language-setting-in-chatgpt. Accessed 2 Aug. 2025.

“Introducing Study Mode.” *OpenAI*, 29 July 2025, openai.com/index/chatgpt-study-mode/. Accessed 2 Aug. 2025.

“Khanmigo.” *Khan Academy*, www.khanmigo.ai/. Accessed 2 Aug. 2025.

Kosmyna, Nataliya, et al. “Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task.” *arXiv*, 2025, arXiv:2506.08872

McInerney, Kristen, and Nicole Mills. “Bridging Language and Learning: Empowering Multilingual Learners in STEM.” *DoD STEM*, U.S. Department of Defense, Sept. 2023, dodstem-assets.dodstem.us/files/DSEC_Literacy%20in%20STEM_MLLs_Toolkit_FINAL.pdf. Accessed 9 Aug. 2025.

Mitchell, Corey. “English-Learners Often Denied Full Access to STEM Education, Report Finds.” *Education Week*, Editorial Projects in Education, 18 Oct. 2018, www.edweek.org/teaching-learning/english-learners-often-denied-full-access-to-stem-education-report-finds/2018/10. Accessed 23 July 2025.

---. “With More Exposure to Science, English-Learners’ Achievement Soared.” *Education Week*, Editorial Projects in Education, 19 Jan. 2017, www.edweek.org/teaching-learning/with-more-exposure-to-science-english-learners-achievement-soared/2017/01. Accessed 23 July 2025.

“New UNESCO Report Calls for Multilingual Education to Unlock Learning and Inclusion.” *UNESCO*, 18 Feb. 2025, www.unesco.org/en/articles/new-unesco-report-calls-multilingual-education-unlock-learning-and-inclusion. Accessed 23 July 2025.

“Privacy and Education Technology.” *Protecting Student Privacy*, U.S. Department of Education, studentprivacy.ed.gov/privacy-and-education-technology. Accessed 9 Aug. 2025.

Reynolds, Lila. “Lack of Non-English Languages in STEM Publications Hurts Diversity.” *Robert R. McCormick School of Engineering and Applied Science*, Northwestern University, 1 Sept. 2021, www.mccormick.northwestern.edu/news/articles/2021/09/lack-of-non-english-languages-in-stem-publications-hurts-diversity/. Accessed 23 July 2025.

Shoemaker DeMio, Paige. “How States and Districts Can Close the Digital Divide To Increase College and Career Readiness.” *Center for American Progress*, 30 July 2024, www.americanprogress.org/article/how-states-and-districts-can-close-the-digital-divide-to-increase-college-and-career-readiness/. Accessed 9 Aug. 2025.

Smit, Jantien, et al. “Inclusive STEM Teaching from a Language Perspective: Teacher Learning in a Professional Development Program.” *European Journal of STEM Education*, vol. 8, no. 1, 6 Sept. 2023, p. 07. *ERIC*, https://doi.org/10.20897/ejsteme/13643. Accessed 9 Aug. 2025.

“Understanding Multilingual Learners.” *WIDA*, University of Wisconsin-Madison, wida.wisc.edu/teach/multilingual-learners. Accessed 9 Aug. 2025.

“When AI Gets It Wrong: Addressing AI Hallucinations and Bias.” *MIT Sloan Teaching & Learning Technologies*, mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/. Accessed 7 Aug. 2025.

"Which Languages Can You Use for Khanmigo Activities?" *Khan Academy*, 2025, support.khanacademy.org/hc/en-us/articles/33322691603469-Which-languages-can-you-use-for-Khanmigo-activities. Accessed 2 Aug. 2025.

Xiang, Yu-Tao, et al. "Timely Research Papers about COVID-19 in China." *The Lancet*, vol. 395, no. 10225, Feb. 2020, pp. 684-85. *PubMed*, https://doi.org/10.1016/s0140-6736(20)30375-5. Accessed 9 Aug. 2025.

"YouLearn." *YouLearn*, www.youlearn.ai/. Accessed 2 Aug. 2025.

YouLearn [@youlearnai]. "YouLearn now supports 21 different languages! This has been one of our most requested feature especially since our recent global growth. Here's the list of all languages that we support:." X, 3 Oct. 2024, x.com/youlearnai/status/1841999857428447237.

Phishing Email Detection Using General-Purpose Large Language Models

By Ethan Barniv and Iris Fu

Abstract

Phishing emails pose a significant cybersecurity threat, often tricking users into giving up sensitive information (1). Although specialized models can accurately detect phishing (2,3,4,5), they typically require complex training and optimization. This study checks whether general-purpose Large Language Models (LLMs) - OpenAI's GPT4o, GPT-4.1, and o4-mini - can effectively detect phishing emails without extensive customization and time-consuming work. We predicted that the models would be able to achieve results on par with models built by other researchers, while reducing work and effort. Tests on a large, labeled dataset (6) showed that after minimal training, these models achieved over 97% accuracy, nearly matching specialized model performance. In addition, there is greater interpretability through this form of classification, as OpenAI's models explain their rationale behind making their decisions.

Introduction and Related Work

Artificial intelligence has grown rapidly over the last few years, and is being used in fields such as healthcare, finance, robotics, and many more (7). One area AI could help in is cybersecurity and specifically phishing. Phishing is a type of cyberattack where attackers attempt to trick individuals into revealing sensitive information. This is typically done by impersonating a trusted source and causing the user to enter confidential data or download malware. A common form of phishing is via email (8).

A problem that many people encounter is the difficulty of differentiating between an actual email from your bank or service provider, and a malicious one (9). Training an AI model to tell the difference between these could and has helped in solving this problem (10). Even before processing the data, the dataset in use must be large and diverse enough to train an AI and check the model on. AI Driven Phishing Detection Systems used a dataset of 1605 emails of which 1191 were real emails and 414 were phishing (2). A dataset of this size could fail to train a model on edge cases and result in lower accuracy. Ideally, a dataset would consist of more emails in order to be more diverse, like in Al-Subaiey et al., where a dataset of approximately 82,500 emails of which 42,891 were malicious and 39,595 were legitimate (3,6). The data used merged six separate datasets resulting in a larger variety. Al-Subaiey et al. then used a data processing pipeline to remove stop words and reduce words to their base form (3).

Building an AI model for this purpose typically involves preparing the data for training. Methods used in past research include Term Frequency Inverse Document Frequency (TF-IDF) (3), which scores a word based on how frequent it is in the email and how rare it is in other documents online. If a word is used frequently in a specific email but is not a common word in other documents, it is scored highly. Word2Vec (3) is an algorithm which turns words into vectors in a multidimensional space in which the closer two vectors are to one another, the closer the words are in meaning. Random Forest (2,3,4) is an algorithm that uses an ensemble of

decision trees to make predictions. It works by combining the predictions of multiple decision trees, each trained on a different part of data, to improve accuracy. Naive Bayes Classifier (2,3,4,5) is a probability-based algorithm that classifies data based on keywords using Bayes' Theorem. Ogundairo and Broklyn also used deep learning models like RNNs (primarily used for sequential data like text or videos) and CNNs (primarily used on spatial data such as images) to identify patterns or common features used in phishing emails (2). An explainable AI (XAI) algorithm was used called Local Interpretable Model-Agnostic Explanations (LIME), which works by changing a small characteristic in a specific email each time and observing which features influenced the model's decision.

While these methods seem to work well for building an AI model, they raise the question of whether building a specific model to perform the task of detecting phishing emails is necessary or would training a large language model (LLM) on similar data be just as effective. The models used in past research seem to be extremely effective at identifying phishing emails (2,3,4,5). With that being said, cleaning up the data, processing it, and using various complex algorithms may take significant effort. If a general consumer LLM, e.g., OpenAI's GPT-4o, could demonstrate performance on par or better than models trained specifically for phishing email detection, then we can significantly reduce development time for niche use cases like phishing email detection.

In this paper, we leverage open-source consumer models, including OpenAI's GPT-4o, GPT-4.1, and o4-mini, to detect phishing emails, and compare their results to the model built by Al-Subaiey et al. for that purpose. Our results show that general-purpose LLMs can achieve competitive performance on phishing detection tasks. This suggests that using LLMs can reduce the effort required to build task-specific models and offer an alternative for phishing email detection.

Methodology

To test how OpenAI's GPT 4o, GPT 4.1 and o4-mini models compare with Al-Subaiey et al.'s model, we utilize the same dataset as the one trained on and tested by them.
Then, we interface with OpenAI's models through their online interface at openai.com, and leverage evaluation metrics such as accuracy, precision, recall and f1 score to compare our results directly with those found in Al-Subaiey et al.'s paper.

The data used in this study was created by Al-Subaiey et al., it is the same one used in their paper "Novel interpretable and robust web-based AI platform for phishing email detection" and it is made publicly available on Kaggle (6). The dataset is a combination of six separate sources (Enron, Ling, CEAS, Nazario, Nigerian, SpamAssassin) and contains approximately 82,500 labeled emails, of which 42,891 are phishing emails, and 39,595 are legitimate. Leveraging this pre-existing dataset, a new file, emailsNoLabel.csv, was created that stripped the labels column and retained only the emails themselves. This was then used in a later step to feed into OpenAI's ChatGPT models.

Baseline Test

For the baseline, emailsNoLabel.csv was inputted to ChatGPT-4o. The model, up to this point, received no training. The goal was to observe how this model would perform with no training, compared with later models that were tested with training. The input prompt was:

> *"This is a file containing emails. Some of the emails are phishing emails and some of them are legitimate. Can you please create a new file by the name of "answer_ChatGPT" that is your predictions on which emails are phishing it should have 2 columns - one is the emails in the file I just uploaded, the other full of 0's and 1's where 1 means phishing and 0 means legitimate. Please predict which are which in whatever methods you think will achieve the highest accuracy*"

In-Context Training

Next, to measure the effectiveness of training GPT-4o, o4-mini, and GPT-4.1, 20 percent of the data was sampled. The data was sampled using ChatGPT o4-mini-high so that the sample contained 16,497 emails, of which 8,578 emails were phishing and 7,919 were legitimate, maintaining the original dataset's ratio between phishing and legitimate emails. The sample was then examined through a manual sample-based verification process to ensure that ChatGPT did not generate non-existent emails. Three models (GPT-4o, o4-mini, GPT-4.1) were then trained on the sample using the prompt:

> *"I have attached two files. Both files contain emails - some of the emails are phishing and some are legitimate. Please train yourself with the data in the file named emails 20 percent sample.csv. In it are emails that are labeled - 1 representing a phishing email and 0 representing a legitimate email. Afterwards, generate a new file called predictions [model name].csv available for download containing all of the emails in emails no label.csv where one column contains the emails in emails no label.csv and in another column are your predictions for which emails are phishing and which aren't - where 1 3 represents phishing and 0 represents legitimate. Try to achieve the highest possible accuracy, precision, f1 score and other metrics."*

This prompt was inputted to all three models to ensure consistency across experiments. Each model then generated a file containing two columns, as specified in the prompt.

Evaluation Metrics

To evaluate each model's success, the following metrics were used:

- True Positives (TP): Correctly predicted phishing emails.
- True Negatives (TN): Correctly predicted legitimate emails.
- False Positives (FP): Legitimate emails incorrectly labeled as phishing.
- False Negatives (FN): Phishing emails incorrectly labeled as legitimate.

Accuracy:

The proportion of correct predictions (both true positives and true negatives) out of all predictions made.

Accuracy = $\frac{TP + TN}{TP + TN + FP + FN}$

Precision:

The proportion of true positive predictions among all positive predictions made by the model.

Precision = $\frac{TP}{TP + FP}$

Recall:

The proportion of actual positive cases that were correctly identified by the model. It is also known as sensitivity or true positive rate.

Recall = $\frac{TP}{TP + FN}$

F1 Score:

The combination of precision and recall. It provides a balance between the two metrics, especially when there is an imbalance in class distribution.

F1 score = $2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$

False Positive Rate (FPR):

The proportion of actual negative cases that were incorrectly classified as positive.

FPR = $\frac{FP}{FP + TN}$

False Negative Rate (FNR):

The proportion of actual positive cases that were incorrectly classified as negative.

FNR = $\frac{FN}{FN + TP}$

Specificity**:**

The proportion of actual negative cases that are correctly identified as negative.

Specificity = $\frac{TN}{TN + FP}$

Ensuring Isolation of Data

During the whole conducting of this study, ChatGPT's "Reference saved memories" feature was turned off, and each model's training and file generation was done in a separate chat. This was a necessary step to ensure that models were generating the results on their own, and not by referring to labels already provided to another chat. When asked, "Are you able to use data from my other chats?" the response given was:

> *"No, I can't access data or files from your other chats. Each conversation is stateless and isolated for your privacy, unless you explicitly upload a file or share details again here. If you need to refer to something from a previous chat, just re-upload the file or paste the relevant content, and I'll help you from there."*

While this response may not give us 100 percent certainty that data is not shared between chats, it certainly gives us assurance, combined with OpenAI's documentation explaining the feature of preventing information from being shared between chats.

Results

Three models were tested (GPT-4o, o4-mini, GPT-4.1). In addition, GPT-4o was tested without training on the 20 percent sample. This was done to check whether general-purpose LLMs can perform on par with task-specific models.

Each model was trained on 20% of the dataset (6) and then asked to predict which emails were phishing and which were legitimate. In addition, one of the models was evaluated without any training in order to learn whether or not it is a necessary step for detecting phishing emails using LLMs.

Model	Accuracy	Precision	Recall	F1-Score	FPR	FNR	Specificity
o4-mini	98.13%	97.56%	98.76%	98.15%	2.64%	1.23%	97.36%
GPT-4.1	97.78%	97.75%	97.99%	97.87%	2.45%	2.41%	97.55%
GPT-4o	97.88%	97.56%	98.42%	97.98%	2.69%	1.59%	97.31%
GPT-4o (No Training)	52.71%	57.61%	39.17%	46.63%	32.18%	60.83 %	67.82%
Al-Subaieyet al.'s	**99.19%**	**99.00%**	**99.00%**	**99.00%**			

Table 1. Performance of each model according to metrics in comparison to a model built especially for phishing email detection.

The best performing model, o4-mini, in this experiment almost reached the results attained by Al Subaiey et al. Further, GPT-4o with no training significantly underperformed those results (Table 1).

While the quantitative results showed slight underperformance (Table 1), it is interesting to note the boost in interpretability that is obtained via this experimental method. Due to the ability to interact directly with the model and ask follow-up questions, it is possible to get a sense of the reasoning behind the model's predictions.

An example interaction showing model interpretability is shown below.

Figure 1. Example interaction between user and ChatGPT model to show interpretability:
Prompt: In "4.1 predictions.csv" there is a mistake in line 1599 starting "new email yo" ... Why do you think that you predicted it to be phishing when in reality it is labeled as legitimate?
Model Response: Why did the model flag this as phishing?

1. Keyword “free email” and “hotmail”: Phishing emails often use “free,” “email,” and free email services like Hotmail, Gmail, Yahoo to create new accounts for scams.

2. Impersonal, informal language: The opening “yo bro” and lack of professionalism can trigger models trained on business correspondence vs. scams/spam.

3. “Get private free email”: Many phishing messages urge recipients to click on a link or get a “private/free” account. This phrasing is highly suspicious for most models.

4. Presence of an (obfuscated) URL: Even though “http www hotmail com” isn’t clickable, URLs—especially to major services—are a phishing hallmark.

5. Short, low-content message: The message is brief, offers little context, and is primarily about changing contact information—again, a common social engineering pattern.

Discussion

The significant difference in performance between GPT-4o with no training and the models trained on 20% of the labeled dataset (6) highlights the importance of training - even when using a general consumer LLM. Specifically, GPT-4o without training achieved only 52.71% accuracy, and notably lower precision (57.61%) and recall (39.17%) compared to trained models, which consistently showed results above 97% in all metrics. We think that this disparity in the results can be attributed to the model’s unfamiliarity with characteristics of emails, as well as the terminologies “phishing” and “legitimate” when used in this context. Without training the model before asking it to classify the emails, GPT-4o had no familiarity with similar tasks, and as a result, its predictions were close to just random guesses.

Giving the models 20% of the labeled dataset (6) for training improved the models’ performance. This showed the models’ ability to learn new tasks they had no experience with beforehand. Further, just training on data a single time allowed the models to recognize patterns and achieve better results. 4o-mini slightly outperformed GPT-4o and GPT-4.1 in identifying which emails were phishing.

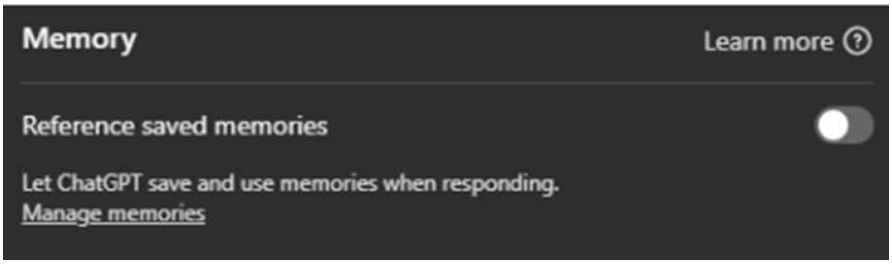
GPT-4-mini is particularly good at complex, multi-step tasks because of its advanced reasoning capabilities (11). The difference in the results between the models was marginal (normally around 0.5%), which could mean that the difference is a result of coincidence. However, o4-mini’s better performance is probably because of its advanced reasoning capabilities (11), making it a bit more effective in classification tasks like identifying phishing emails.

Although all three models performed extremely well, they slightly underperformed compared to the model built in the paper by Al-Subaiey et al. (3), which was specifically built for phishing email detection. Their model, which was created using various techniques like TF-IDF, was optimized to succeed at the specific task of phishing detection. However, ChatGPT’s models have several purposes and even when provided with some labeled data, they are not built exactly for email classification.

These results are extremely promising in pointing to the usage of general-purpose LLMs like ChatGPT’s models for specific tasks, but there could be some drawbacks. One major

limitation of this setup is that it is uncertain how ChatGPT manages data across separate chats and users. Although the feature of using data from separate chats was turned off during the entire research (Figure 2), we cannot be completely sure whether data provided in one chat is never used in other chats or with other ChatGPT users. Future research should try and investigate this topic and if needed, find more fail-safe ways to bypass these problems.

Figure 2. ChatGPT's "Reference saved memories" feature turned off.



Another direction for future work could be fine tuning the models to be better at the task of classification (could be of data that is not emails). While this study used a prompt-based approach, more finetuning, especially using the Open AI's API, could lead to better performance. Customizing models using the API could lead to more classification-oriented performance and achieve even better results on classification tasks. The results support Ogundairo and Broklyn's findings that AI driven systems demonstrate the ability to detect phishing attacks, although their methods such as deep learning differ from using an open-source consumer model.

Al-Subaiey et al. suggested in the Future Work section to continue using explainable AI. The method of using consumer models to detect phishing emails excels at interpretability-researchers can prompt the specific LLM on why it made a specific decision, i.e. why it labeled one email as phishing, and another as legitimate. It is demonstrated that consumer model LLMs offer a great solution to using explainable AI in a simple and user-friendly way (Figure 1).

Another suggestion that Al-Subaiey et al. discussed is to make models exhibit greater cross-domain adaptability. This experiment with OpenAI's consumer-facing models directly addresses this suggestion, as models such as GPT-4o are great for general purposes and are adaptable to many use cases. Behind the scenes, it is likely that OpenAI's models leverage transfer learning and that expertise in one domain improves its performance in other areas.

Conclusion

In conclusion, this research explored the effectiveness of general-purpose large language models, specifically OpenAI's GPT-4o, GPT-4.1, and o4-mini, in detecting phishing emails compared to specialized AI models. Initially, GPT-4o without training performed at a level similar to random guessing, showing the importance of providing training to LLMs. Training on 20% of a labeled dataset substantially improved performance across all models, which caused all models to achieve results of 97% and better across all metrics. Notably, o4-mini demonstrated

marginally better performance, which is likely because of its enhanced reasoning capabilities (according to OpenAI's website).

Despite high effectiveness, all three models slightly underperformed compared to the model specifically designed for classifying phishing emails developed by Al-Subaiey et al. However, general-purpose LLMs like OpenAI's models offer significant advantages in effort and interpretability, as shown by their ability to give explanations for specific decisions. This aligns with Al-Subaiey et al.'s recommendation for further exploration of explainable AI.

Future research directions could include fine-tuning these models using APIs to achieve even better performance and diving deeper into how ChatGPT uses the data given by users. Overall, this study highlights the potential of consumer-accessible LLMs as tools in cybersecurity applications, particularly in phishing detection, provided that they are trained.

Works Cited

1. CSRC Content Editor. Phishing - Glossary | CSRC. csrc.nist.gov/glossary/term/phishing.
2. Ogundairo, Obaloluwa and Peter Broklyn. “AI-Driven Phishing Detection Systems.” Journal of Cyber Security, Aug. 2024, www.researchgate.net/publication/382917933.
3. Al-Subaiey, et al. “Novel interpretable and robust web-based AI platform for phishing email detection.” Computers and Electrical Engineering, vol. 12, part A, Dec. 2024, doi.org/10.1016/j.compeleceng.2024.109625.
4. Rawal, Srishti, et al. “Phishing Detection in E-mails Using Machine Learning.” International Journal of Applied Information Systems, vol. 12, no. 7, Oct. 2017, pp. 21–24, doi:10.5120/ijais2017451713.
5. Eze, Chibuike Samuel and Lior Shamir. “Analysis and Prevention of AI-Based Phishing Email Attacks.” Electronics, vol. 13, no. 10, May 2024, p. 1839, doi:10.3390/electronics13101839.
6. Abdullah Alam, Naser. Phishing Email Dataset. Kaggle, 2024, doi:10.34740/kaggle/ds/5074342.
7. “Applications of Artificial Intelligence (AI) | Google Cloud.” Google Cloud, cloud.google.com/discover/ai-applications. Accessed 17 July 2025.
8. “What Is Phishing? - Meaning, Attack Types and More | Proofpoint US.” www.proofpoint.com/us/threat-reference/phishing. Accessed 17 July 2025.
9. “Avoiding Social Engineering and Phishing Attacks.” America’s Cyber Defense Agency, 1 Feb. 2021, www.cisa.gov/news-events/news/avoiding-social-engineering-and-phishing-attacks. Accessed 15 July 2025.
10. “The Role of AI in Email Security: Beyond Phishing Detection.” CSO, 27 Aug. 2024, www.csoonline.com/article/3496828/the-role-of-ai-in-email-security-beyond-phishing-detection.html?. Accessed 16 July 2025.
11. “Models.” OpenAI API, platform.openai.com/docs/models. Accessed 20 May 2025.

Maximizing Franchise Value in the NFL: Actions Owners Can Take to Drive Growth
By Aidan Moses

Abstract

This paper examines factors driving NFL valuation growth and strategies that owners can use to maximize franchise values. Two-thirds of NFL revenue is distributed evenly amongst the 32 teams and comes from national sources such as media contracts and league sponsorships, while the remaining unshared revenue comes from local initiatives such as stadium operations, team-owned real estate, market size, and sponsorships, and is earned unevenly. Case studies of teams like the Cowboys and Rams illustrate the impact of luxury suites and mixed-use developments on team valuations. Although on-field factors generally have minimal long-standing effects, only accounting for 2.6% of valuation growth, superstar player acquisitions or Super Bowl wins can drive spikes in valuation through merchandise sales and fan engagement. This analysis also considers the challenges associated with valuation growth, and addresses the rising disparities between large- and small-market teams. Findings suggest that the most effective way to increase valuation is through business-oriented revenue streams, especially relocating to larger markets or investing in team-owned real estate, rather than focusing on team performance.

Introduction

In recent years, the valuation of NFL teams, or how much they are worth, has skyrocketed, up to as high as $10.1 billion, making them some of the most valuable sports assets in the entire world. For owners, a comprehensive understanding of the factors driving this exponential growth is critical to increase the valuation of their team. However, there are various methods and factors that contribute to increasing team valuation. Since two-thirds of the revenue in the NFL is shared among the 32 teams, it is more important to look at trends with the NFL's unshared revenue streams, such as stadium investment, market size, team-owned real estate, on-field performance, and team sponsorships. Finding out which of these methods is the most lucrative can help owners to understand the steps they need to take to become the most profitable and highest valued in the entire league. This paper examines a variety of these methods, analyzing them against past valuation data to determine which of them is the most effective as a potential guide for NFL owners of the future. In that regard, the question that will be answered through this paper is: What are the most effective things an NFL owner can do to raise the valuation of their team?

Literature Review

NFL Valuation Factors

NFL franchises are worth more than ever before, and their valuations are continuing to rise. NFL team valuations are determined by a combination of shared and unshared revenue streams, largely as a consequence of the NFL's centralized revenue-sharing model, in which

roughly 63% of the revenue from the league gets distributed among the 32 teams. More specifically, this shared revenue comes from a combination of league-wide sponsorships, major media contracts, and gate receipts (Appraisal Economics 2017).

In 2023 alone, the NFL made an impressive $20.5 billion in revenue, significantly surpassing other major sports leagues by over $8 billion (Badenhausen 2024). Of that $20.5 billion, over $13.5 billion was shared equally among the 32 teams, providing each team with over $400 million. Additionally, each team received another $25 million from the NFL's gate receipt distribution system, where 34% of every team's ticket revenue is pooled and then distributed evenly (Badenhausen 2024). This shared revenue makes up an approximately two-thirds portion of a team's calculated valuation.

However, the other third of the league's revenue is earned unevenly and can come from a variety of sources. Some of these unshared revenue streams that contribute to valuation discrepancies include local media deals and stadium use, which may also be impacted by team performance, market size, location, and franchise management. Teams can generate additional unshared revenue from non-NFL events in stadiums, unshared ticket sales, and branding deals. These factors vary significantly by franchise and make up the main reasons as to why some teams are valued so much higher than others (Moore 2023). Of these factors, some have a greater effect on valuation than others. For example, in 2023, $3 billion of the $20.5 billion of total NFL revenue came from luxury suites and ticket sales, or about 15% of the total (Badenhausen 2024). This $3 billion was not shared and remained with the individual teams. After removing the 15% of revenue from luxury suites and ticket sales and the 66% from shared revenue, this leaves only 19% of the total revenue attributable to other unshared, localized sources.

The team with the highest valuation in every single year is the Dallas Cowboys. In 2022, their valuation was over $8 billion and they received over $1 billion in revenue (Forbes 2024). This higher valuation is attributed to their strong market presence, with Dallas Fort-Worth having the 4th biggest market in the United States, a stadium size expandable up to 105,000, robust branding, and non-NFL events, such as the 2021 Rose Bowl and the 2014 NCAA men's final four (Forbes 2024). After the Dallas Cowboys, the other top-five franchises in valuation that year were the Patriots, Giants, Rams, and Commanders, respectively (Moore 2023). Analyzing the characteristics of the teams that are valued the highest can offer insights into the biggest potential drivers of franchise valuation change.

Unshared Revenue

Stadiums, Locations, Markets

The NFL operates with a two-part business model: national revenue and local revenue. While national revenue is shared among the 32 teams, the local revenue can help viewers comprehend why a team like the Patriots is worth $7.4 billion while a team like the Jaguars is only worth $4.6 billion. However, it is important to acknowledge that all of the unshared revenue

streams are generally smaller compared to the shared revenue solely due to the substantial national TV deals and league-wide sponsorships (Eckstein 2024).

One of the foremost ways that teams can stand out in their valuation is through their stadiums, considering factors such as ownership, size, age, and location. The stadium is often considered the most valuable asset that an NFL team owns, with most generated revenue being retained by the owning franchise (Badenhausen 2024). Teams can increase their revenue by investing in luxury suites. All revenue from these suites goes directly to the team and because they command high prices, they are heavy revenue builders for teams. One example of a team who invests in luxury suites is the Cowboys, who make over $130 million annually from them (Rascher 2004; Badenhausen 2024). A disparity is evident when comparing the Raiders, who make an average of $75,000 per game per luxury suite to the Jaguars, who only make $16,000 per luxury suite, significantly contributing to their difference in valuation (Bisson 2025).

Teams can also enhance profit by increasing the stadium capacity (if the footprint allows) and investing in more amenities for fans. The average stadium size in the NFL is about 70,000 seats, and since games usually sell out or are near capacity, these tickets make a consistent amount of revenue, even if roughly one-third of it is shared through the gate receipt system. Despite high upfront costs, renovating a stadium in the NFL should pay off monetarily. This is because of the high demand for the limited number of home games, which is typically eight or nine home games per season. Renovating a stadium also allows for higher revenue during non-NFL events. Moreover, about 1/3 of funding for stadiums is public through subsidies or tax breaks, meaning that NFL teams can reap rewards without taking out all of the financial burden (Eckstein 2024; Klawans 2024).

Additional aspects of a stadium that account for the remaining unshared revenue are outright ownership and non-NFL events. Because stadiums represent significant assets for franchises, not owning the stadium would result in lots of missed yearly revenue. This is evident in the example of the Chargers versus the Rams, where the Rams are valued at more than $2 billion above the Chargers. Stan Kroenke, the Rams owner, financed the stadium and thus the Rams own 85% of the stadium compared to the Chargers' mere 15%, meaning that the revenue from non-NFL events is more lopsided (NBC DFW 2024). Another example of a team that completely owns their stadium versus a team that does not is the Patriots versus the Bills. The Patriots have complete ownership of their stadium and now have a valuation of over $7.4 billion. The Bills, on the other hand, do not own their stadium and have a valuation of $4.2 billion, a big difference from the Patriots. On a league-wide basis the average valuation of teams with stadium ownership is $6.4 billion, while the average valuation of teams without ownership is $5.54 billion. This discrepancy can be attributed to control over stadium revenue and increased non-NFL event revenue. These examples show that ownership percentage in stadiums can tie directly into a team's unshared revenue, which positively impacts long-term valuation.

Non-NFL events are the last source of revenue that teams earn in their stadium, as the team that owns the stadium receives revenue from all events held there, regardless of whether the event is NFL related or not. This means that teams in busy metropolitan areas that host many

events can generate substantial supplemental revenue during the offseason. For example, the Raiders hosted over 15 major non-NFL events in Allegiant Stadium in 2024, guaranteeing them another revenue stream, while the Jaguars only hosted about two such events (Ross and Woodworth n.d).

Beyond the stadium, the location and the local market of the team are equally as important to valuation. Teams located in the most populous cities and metro areas are generally more successful in obtaining high value sponsorships, attracting and retaining viewers and fans at games, and gaining overall support. As of 2023, four out of the top five teams in terms of media viewership were in the top half of market size (Beach 2024). However, it is important to note that teams with established and loyal fanbases like the Packers and Bills are also able to have significant media presence. Generally, though, teams such as the New York Giants or the Los Angeles Rams that are located in some of the most populated metro areas in the world are extremely valuable. They have some of the largest markets available to them, and teams that have access to these markets tend to utilize them and earn the most revenue (Moore 2023).

Team-Owned Real Estate

Increasingly, sports facilities are being used as central anchors for mixed-use developments. As of 2024, across American major league sports, there were 43 real-estate developments attached to stadiums with more under construction (Talkington 2024). This essentially means that owners are building real-estate developments around their stadium to increase revenue throughout the year. This approach is particularly attractive because many teams have already maximized most stadium-based revenue streams, such as by adding luxury suites and expanding the number of seats. The other benefit of team-owned real estate is that the revenue that comes from these developments is year round, meaning that it is not as dependent on the limited number of NFL games. This revenue is also unshared, meaning that it directly benefits the team (Harrison 2018; Mueller 2024).

One way that teams can leverage this real-estate opportunity is by creating mixed-use developments surrounding or nearby their stadium. For example, the Patriots built Patriot Place adjacent to Gillette Stadium in 2007, which contains a hotel, theater, stores, and other amenities. This area generates consistent revenue for them over the year, and receives over 9 million visitors annually compared to the only 500,000 visitors that enter the stadium. This real estate is a primary driver of the Patriots' increasing valuation (Harrison 2018; Dart 2023). About nine other teams are becoming year-round commercial entities through real estate as well. Some of these teams include the Packers with Titletown, Rams with Hollywood Park, the Vikings with Viking Lakes, and the Cowboys with the Star in Frisco. All of these teams that have this type of property tend to be higher in valuation than those who do not, with an average of $7.14 billion for those who have real estate, versus only $5.69 billion for those who do not, reflecting the value for teams of owning team real estate. (Harrison 2018; Dart 2023).

In today's NFL, teams that control and monetize the real estate around their stadiums are evolving into year-round commercial enterprises—redefining what it means to be a sports

franchise. By investing in team-owned real estate, owners transition into roles of real-estate investors and even city planners as well. After moving to Los Angeles, Stan Kroenke, the Rams owner, helped develop Hollywood Park, a 298 acre area around the stadium consisting of housing, a theatre, and other amenities. This area has helped the Rams to remain competitive financially amid rising operational and real-estate costs while also increasing their valuation over time (Dart 2023). Because of this increased infrastructure, the Rams' properties are being used 365 days a year rather than just for eight or nine home games (Dart 2023).

Real-estate ownership can also have side effects for teams. On average, NFL stadiums host around five major events yearly beyond NFL games, but now, SoFi stadium, owned by Kroenke Sports and Entertainment, hosts more than 24 events due to its appeal. Although the teams do not retain all the profit, these events create millions of dollars in extra revenue. Even more importantly, by adding infrastructure and mixed-use developments, teams are able to increase the desirability of their location (Dart 2023). For example, one of the main reasons why the Packers were able to host the NFL draft in 2025 is because of their building of the Titletown real-estate development. This real estate increased the appeal of Green Bay and made it more of a popular destination for other events. In summation, due to their investment in real estate surrounding the stadium, the Packers reaped the monetary benefits (Dart 2023).

Local Sponsorships and Naming Rights

One of the easiest ways for NFL owners to get a steady stream of revenue is by selling a stadium naming rights deal. In these deals, a company who is looking for more brand recognition can pay an NFL team to have rights to name their stadium. For example, the 49ers signed a $170 million deal with Levi Strauss & Co for their stadium naming rights in 2024 (Eckstein 2024). This has provided them with steady revenue as the deals generally last for a set amount of years. Because of the high demand and popularity of the NFL, corporations often want to be associated with these large franchises, motivating them to invest significantly in naming rights agreements (Eckstein 2024).

Another way that teams can acquire revenue locally is by selling sponsorships that are team and location specific. These sponsorships operate independently of league-wide sponsorships, and all revenue directly benefits the team. One of the teams that has been extremely successful with these types of local sponsorships are the Dallas Cowboys. Knowing that Visa and Coca-Cola were two league-wide sponsors, Jerry Jones, their owner, forged sponsorship deals with two of their biggest rivals, American Express and Pepsi (NBC DFW 2024). These two massive sponsorship deals have provided the Cowboys with consistent revenue for many years.

However, this method of earning revenue may be much more difficult than some of the others, especially ones of the same size as Pepsi and American Express for the Cowboys. Most local sponsorships generate less revenue for the teams than national brand deals, and because of this provide a smaller revenue stream for the team. Nevertheless, local sponsorships can still be

an effective method to gain more revenue and thus increase valuation over time (NBC DFC 2024).

On-Field Factors Affecting Valuation

Performance

As a whole, the valuation of NFL teams tends to trend upward regardless of on-field success. Between 2012-2024, the valuations of NFL teams have skyrocketed an average of 384% per team, even the teams with poor performance records. However, winning percentage may have a slight effect on valuation. When looking at the relationship between a team's winning percentage during this period and the resulting valuation percent increase, winning percentage only seemingly accounts for about 2.6% of the change in valuation, which is a relatively small number compared to other examined factors (Tamir 2022). For owners, this suggests that investing in non-NFL facilities, real estate, or other diversified revenue streams may yield higher returns than focusing on team building, as teams have grown in revenue and valuation even with poor performance.

Paradoxically, making the playoffs for an NFL team also provides almost no financial gain for owners. This is because the league centralizes 100% of the playoff ticket revenue and distributes it evenly among the 32 teams. This is vastly different from the comparatively small 1/3 of ticket revenue that is taken during the regular season. The only benefit that teams have for making the playoffs financially is that they retain concessions and parking money, which are notably smaller than ticket sales (Badenhausen 2022).

Teams can even incur financial deficits during the playoffs due to player bonuses. For example, the Buccaneers paid Tom Brady $2.25 million for winning the Super Bowl as a part of a contract bonus. Players also receive bonuses for being in the playoffs, with $150,000 going to each player on the winning team during the Super Bowl, further increasing the yearly payout from the team (Badenhausen 2022).

Overall, the entire NFL benefits from the playoffs because of the large sums of money from TV broadcasts, ticket sales, and other revenue streams. The playoffs can be beneficial for individual teams on occasion, but typically only if the team wins the Super Bowl. This is because winning a Super Bowl raises fan excitement and engagement, increases merchandise sales, and likely increases season ticket sales (Badenhausen 2022).

The Star Player Effect on Valuation

While team performance only has a modest effect on valuation, star players can drive meaningful spikes in revenue. In general, superstar athletes, no matter the sport, can bring excitement and success to any sports franchise. In the NFL, though rare, the arrival or presence of a superstar player can be seen either monetarily or through team success (Liu 2012).

One example of this phenomenon is with Patrick Mahomes in Kansas City. His dramatic rise in popularity and personal on-field success coincided with an increase in fans per home

game, up to 75,973 fans in 2018, which was Patrick Mahomes' first starting season. His rise also led to a surge in merchandise sales and a higher rise in ticket prices, an increase of 82% compared to the NFL average of 39% (Koebert 2024). Furthermore, Patrick Mahomes' presence in Kansas City has led to more playoff appearances and better records, ultimately elevating the team's popularity and thus increasing revenue (Dent 2023).

Another example of superstar movement in the NFL is Tom Brady. Even after leading the Buccaneers to a Super Bowl win in his first year with the team, when he decided to return to Tampa Bay for a second season, the entire local economy was boosted. Whether it was ticket sales or merchandise, all of the excitement and revenue that came from Brady's return went directly to the Buccaneers as an organization, as it was not subject to revenue sharing. That year, Brady's jersey was the league's top seller, games saw higher attendance, and the overall attention paid to Tampa Bay increased (Cridlin 2022). All of this success is largely attributable to Tom Brady's arrival and immediate success, signifying the impact that a superstar player can have for an NFL team.

However, there are some caveats that come with adding a star player to a new team. In the NFL, acquiring a player of the caliber of someone like Brady or Mahomes is extremely rare, making it extremely difficult to find a star player who will greatly increase the valuation of a franchise. In the NBA, for example, superstar movement is often much more impactful, due to smaller roster sizes and the nature of player movement in the NBA. This explains why players like Lebron James or Kevin Durant had such big impacts on their new and old teams when they changed teams (Liu 2012; Cridlin 2022). In the NFL, however, only a small number of high-caliber players can have this impact, making it an unreliable way to increase revenue. Furthermore, the media often overstates the actual extent of players' impact on team valuation and revenue (Dent 2023).

Team Building: Stars vs Supporting Cast vs Popular Players

This section focuses on the best way to ensure a team is successful monetarily. Although on-field performance only accounts for a small percentage of influence on revenue, it is important to understand the best way to increase fan interest to maximize revenue, whether that is by prioritizing winning or by trying to acquire the right players for their goals (popularity or success). There are a few different options to try and achieve this goal.

The first strategy is acquiring as many star players as possible, often termed "star stacking." To summarize, adding a high number of star players does not necessarily result in better performance or more money, based on data collected, and ensuring to balance team-building strategy with high-end talent is essential to achieving optimal success on the field (Loignon 2024). While having a lot of star players can undeniably attract attention to the team, it is more effective to develop players collectively to become a better unit. Furthermore, adding more stars often yields diminishing returns, and tends to have little to no correlation with a better record. It is generally more beneficial to have one or two high-impact, fan-favorite stars than to overload the roster with the maximum number of stars possible (Loignon 2024).

The second approach emphasizes building a well-rounded team. While franchise quarterbacks are extremely desirable in the NFL due to their scarcity, having a franchise quarterback assures the team almost nothing in terms of performance, especially if most of the salary cap is allocated to the quarterback without investing a lot in a supporting cast. In 2022, the AFC Championship had quarterback led teams, including the Chiefs and Bills, but the NFC Championship had the most well-rounded teams, with the Eagles and the 49ers (White 2022). Although both the Eagles and the 49ers had solid QBs, they had an even better foundation and supporting cast around the QB, allowing them to succeed. Conversely, teams such as the Broncos and Browns, despite acquiring high-profile quarterbacks, experienced significant financial loss due to their unexpected lack of on-field success. This suggests that it may be more beneficial for team success to distribute the salary cap throughout the entire team instead of overpaying for a single position (White 2022).

The final strategy focuses on strategically acquiring players who are fan-favorites or commercially popular. For the 2024-2025 season, Saquon Barkley topped the league in merchandise sales, followed by Patrick Mahomes, Jalen Hurts, and Jayden Daniels, respectively. A lot of the younger players, like JJ McCarthy and Jayden Daniels, are projected to be at the top of sales next year (NFLPA 2025). Since jersey sales revenue is unshared, these spikes directly boost team income, especially if timed with playoff appearances or key contract years. Instead of overpaying for a top player at a position, it can often be more beneficial to have strong players on their rookie deal or an early deal who generate a lot of revenue from merchandise sales. This method allows the team to succeed monetarily while also being competitive in gameplay (NFLPA 2025).

Challenges in Valuation

Despite how much teams' valuations have risen in recent years, there are still some problems regarding valuation that create league-wide tensions and should be addressed. The main challenges with valuation arise with market disparity between different teams league wide. The current system gives big-market teams an advantage in earning potential over small-market teams, as the bigger market teams have larger fanbases, better luxury stadiums, and often more valuable sponsorships. Big-market teams may generate over $100 million annually from these sponsorships, while the smaller market teams do not reach close to that benchmark (SponsorUnited 2025). Additionally, teams located in more populated cities hold more events than smaller cities, further augmenting their unshared revenue streams (Rascher 2004; Ross and Woodworth n.d). In short, this gives them the potential to earn more money to work with than some of the smaller teams.

With this increased revenue potential, the bigger market teams often have better facilities, fund better training and scouting operations, and pay their staff more money. This can create a competitive imbalance where wealthier teams can strive and keep paying more while the smaller market teams have tighter constraints. There are concerns about the creation of a competitive

advantage for the bigger market teams, and some of the possible solutions to the problem can be a greater revenue-sharing plan or more funding for the smaller market teams (Rascher 2004).

There are also concerns about revenue distributions within the league. Although the league generated a new high in revenue of $23 billion in 2024, owners are becoming unhappy over the increasing salary cap and the rising costs of paying players. They claim that the current Collective Bargaining Agreement, or CBA, is unsatisfactory and that increasing costs of players, stadiums, and operation are getting too high to manage (Johnson 2025).

Under the current CBA, players receive about 48% of the league revenue, and owners want to reduce this number as well as stop the ever-increasing salary cap, which has increased 130% since 2011. Owners are also unhappy with rising demands to fund and renovate new stadiums. This is a main reason why NFL owners want an 18th game, because it adds billions of dollars of revenue to the NFL from a single extra contest (Johnson 2025).

Exploratory Data Analysis

This section uses exploratory data analysis (EDA) of NFL team valuations from 2013 to 2024 to examine key factors that contribute to changes in franchise value. Some of the investigations include league-wide valuation trends, year-over-year growth, the influence of metropolitan populations, and more.

Overview of Franchise Valuations

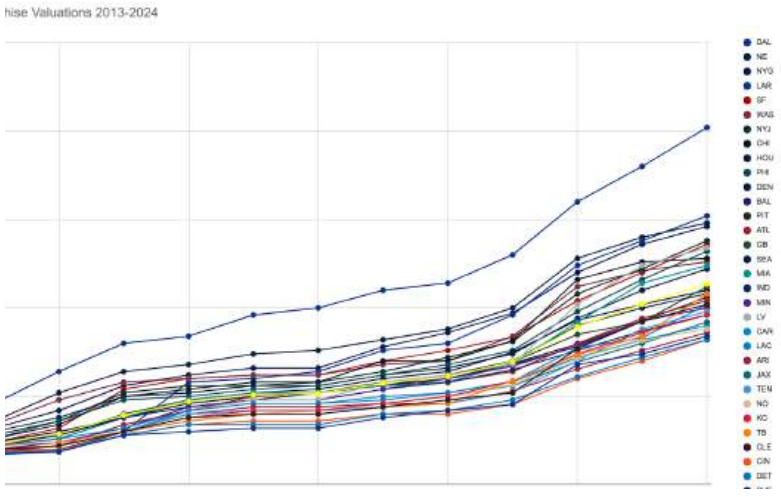


Figure 1
NFL franchise valuation trends since 2013

Figure 1 displays the NFL Franchise valuations of every team in the NFL from 2013 to 2024. Overall, the graph showcases that team valuations have steadily increased during this timeframe, with most teams more than doubling in valuation. This trend reflects the rising value of media rights deals, brand expansion, and general revenue growth due to price inflation. Some teams rank consistently at the top, like the Cowboys, while some of the smaller market teams

like the Jaguars remain consistently toward the bottom. That being said, even the lower-tier teams with less growth compared to the other NFL teams have still drastically increased in valuation. This is a testament to the high and increasing revenue streams that the NFL brings in through the revenue-sharing model. As shared revenue continues to increase, the valuation of every team in the NFL will continue to rise, regardless of the unshared revenue that each individual team makes. However, the analysis also reveals that unshared revenue generation, such as sponsorships, media, and real estate, are crucial for exceptional valuation surges, as exemplified by the Los Angeles Rams' significant jump over other teams, as shown in figure 3.

Yearly Revenue Growth Trends

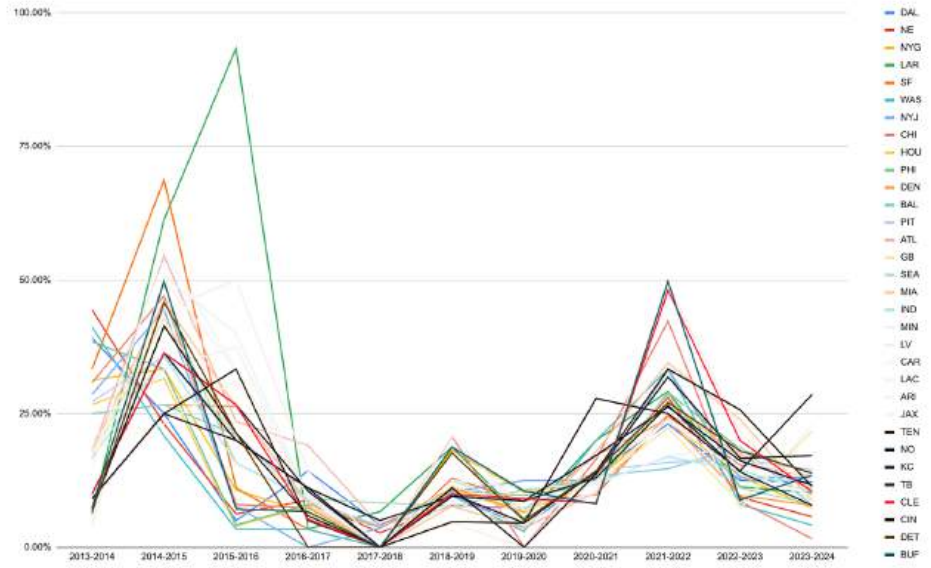


Figure 2
Year-over-year percent change in NFL team valuation

Figure 2 depicts the year-over-year percent change in revenue for each NFL team. While the overall trend is positive, there are two prevalent spikes of growth that arise in 2015 and 2021. These drastic spikes in growth are due to league-wide changes to revenue such as new media deals, which is a primary cause of the 2021 spike, and also franchise sales or post-recession market recovery. In contrast, there is also a significant deceleration in growth observed in 2020, which is attributable to the Covid-19 Pandemic. During this period, many people were unable to go to stadiums to watch the games, decreasing ticket revenue, and also the broader economy was struggling. Examining individual team spikes and dips during these given years can reveal league-wide trends that provide insight into optimal strategies for maximizing specific revenue streams. To do that, each division was in more detail to try and denote which changes in revenue stream had the most drastic impact on the percent change graph.

Divisional Valuation Growth Study: NFC West and AFC West

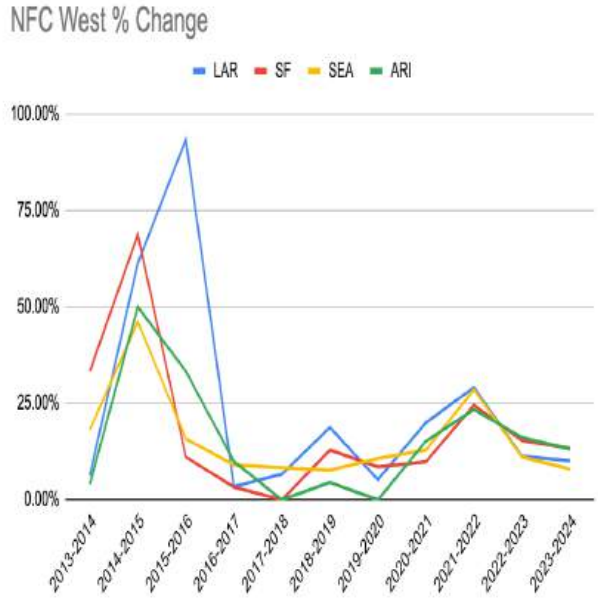


Figure 3
Year-over-year percent change in the NFC West franchise valuations

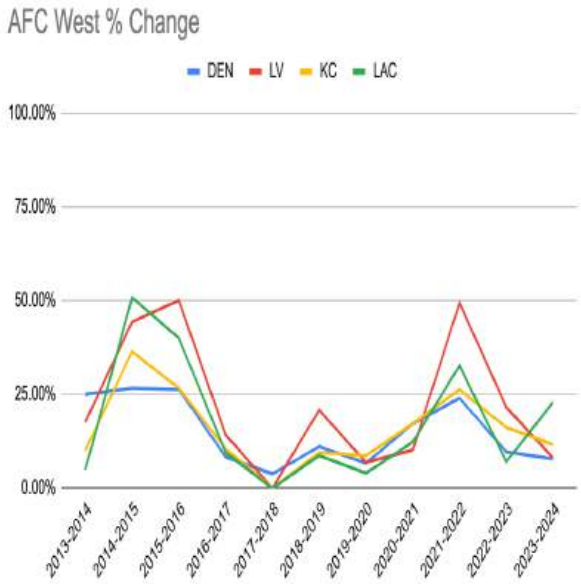


Figure 4 Year-over-year percent change in the AFC West franchise valuations

Figure 3 and Figure 4 highlight the percent change year-over-year of the valuations of franchises in the NFC West and the AFC West, two divisions that were featured in Figure 2. Each of these divisions have specific teams who had extremely high relative spikes or dips in valuation in specific years.

One example of significant valuation growth is the Los Angeles Rams from 2015-2016, as seen in figure 3. This growth can be attributed to the Rams' decision to move to Los Angeles for the 2016 season. With this move came the prospect of a new, stronger market with a bigger metropolitan area. This meant that the Rams could unlock greater branding opportunity, fan engagement, and sponsorship through their new market. The Los Angeles Chargers experienced a similar, although smaller, growth spurt during this year, as they moved to the same market as the Los Angeles Rams. The Chargers' lesser increase can be attributed to factors such as the relative proximity from San Diego to Los Angeles, and their initial status as a stadium tenant rather than a stadium owner like the Rams.

Another example of this phenomenon on these graphs is the Raiders moving to Las Vegas. However, their spikes from 2015-2016 and 2021-2022, as seen in figure 4, are more due to the anticipation of potential relocation. These promises of relocation raised excitement and increased fanbase engagement, because there was a larger market in Las Vegas. The latter spike was a direct result of the actual relocation to the larger market. Both the Rams and the Raiders are owners of some of the highest year-over-year change in the entire NFL, with the Rams having an extraordinary 93% growth rate from 2015-2016 and the Raiders having a 50% growth rate from 2015-2016 and 2021-2022. The Chargers also had a greater than 40% growth rate in valuation from 2014-2016. These cases illustrate how relocating to a larger city can significantly

boost a team's valuation, as the new opportunities available outdo any revenue streams that were in place previously.

However, moving cities is not the only effective method that these teams used to have spikes in valuation. One example on Figure 3 is the 49ers from 2014-2015. During this period, the 49ers completed the construction of Levi's Stadium and moved there from Candlestick Park. This new stadium offered more amenities, had a bigger stadium capacity, and contained more luxury suites than Candlestick Park. This change allowed the 49ers to bring in much more stadium based revenue, contributing significantly to their 75% growth in their team valuation, showcasing how a stadium upgrade can considerably boost a team's valuation. Along with their respective city changes, both the Rams and the Raiders also moved into new stadiums in 2020, with the Rams opening SoFi stadium and the Raiders building Allegiant Stadium. The Rams also built a large real-estate complex that year, Hollywood Park in Inglewood, CA, which further expanded their revenue potential. These large-scale infrastructure projects, paired with strategic relocation, explain why the Raiders and the Rams experienced some of the largest valuation jumps in the league during that time.

The last trend that can be seen in these two graphs is the relative insignificance of team performance and player acquisition when compared with teams that are investing directly into business infrastructure. For instance, the Chiefs have not had a losing season on the timeline of the graph, and yet their growth is not as great as some of the other teams. This may be attributed to the fact that they made few major off-field investments, such as stadium renovation, which meant they had limited opportunity for revenue jumps. However, their growth was still solidly upward because they were so successful on the field and won multiple Super Bowls. Similarly, the Denver Broncos acquired Russell Wilson in 2022, which was supposed to change the franchise. However, after that time period there is not any growth on the graph for the Broncos, indicating that excitement died out very quickly. These examples suggest that while team success is beneficial for league-shared revenue, it does not have that great of an impact for individual teams valuation.

Market Size and Valuation

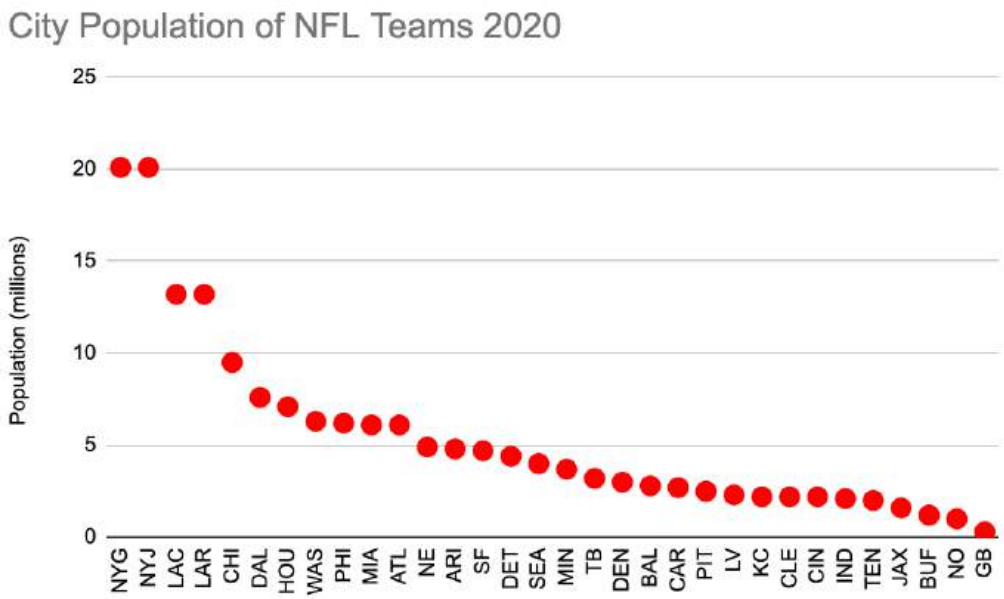


Figure 5
Relative size of the metropolitan area that each team in the NFL is located in

Figure 5 illustrates the relative size in millions of the metropolitan area that every team in the NFL is located in. Teams located in large metropolitan areas, such as Giants, Rams, and Bears, consistently rank among the top 10 in league valuation due to expansive opportunities for marketing, sponsorship, and engagement. Teams that currently reside in some of the smaller markets, like the Saints, Bills, or Jaguars, all occupy the bottom 10 in terms of valuation. If nothing else, this comparison demonstrates the extreme effect that location has on revenue generation and thus valuation. In the future, teams like the Bills or Jaguars should consider relocating to try to tap into a larger market, as they will most likely never be able to lead the league in valuation in their current market. The Packers, however, serve as a compelling counterexample. The reasons why the Packers, the team that resides in the smallest metropolitan area in the league, is still in the top half of valuation is that they have deep-rooted fan loyalty, a historic brand, and they also have invested heavily in Lambeau Field and Titletown so that the team can make solid revenue year round. The Packers' success proves that by strategically investing into infrastructure, teams can compete, at least in part, while residing in a smaller market.

Player Spending and Player Contracts

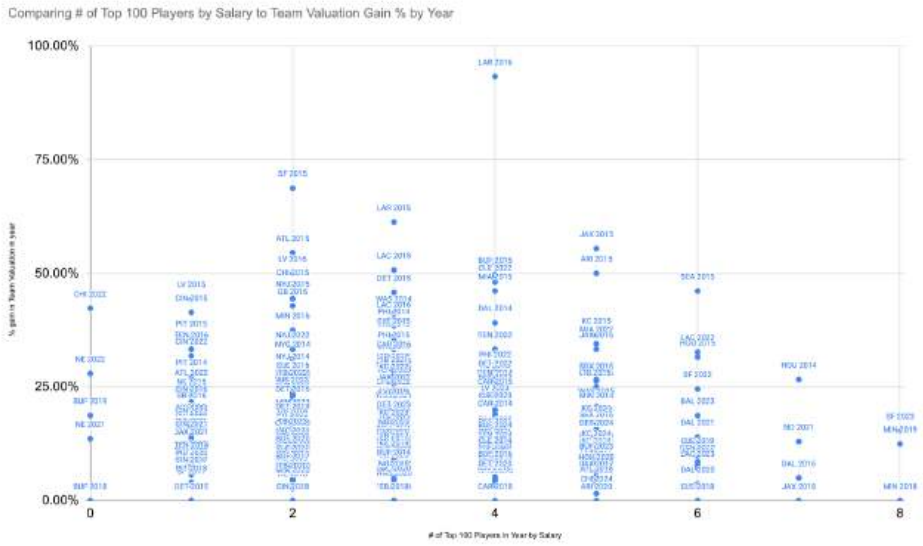


Figure 6
Comparing the relationship between top paid players in the NFL to team valuation increase

Figure 6 explores the relationship between the number of top 100 contracts in current season guarantees that a team has relative to their year-over-year valuation growth. Looking at this graph can demonstrate the relationship between team building trends and overall monetary success. As a whole, the trend line of the graph seems to be in a non-linear "^" shape, with teams in the middle range of player spending achieving the most success. On the graph, teams like San Francisco in 2023 and Buffalo in 2018 have eight and zero top 100 contracts, respectively, that year. Each of these teams saw below average growth that year as well, demonstrating the negative results of being too far from the median. In contrast, the 49ers in 2015, the Rams in 2016, and the Jaguars in 2015 all had between two and five top 100 contracts, or somewhere in the middle on the graph, and they experienced strong valuation increases, reflecting a trend toward trying to stay balanced and even more frugal when it comes to paying players. These examples suggest that both over-spending and under-spending can limit financial gains. However, it is possible for any team, even if they are balanced when paying players, to be extremely poor in valuation growth, and this can be a consequence of broader NFL trends or individual fallouts.

Team Budgeting

When looking at teams in recent NFL history that have spent big on stars, it is revealing to examine how their valuation pans out based on the success of that star player. First, it is significant to look at teams who invested heavily in a given player, but their investment did not pay off. The first example of this is the Broncos. They signed Russell Wilson to a $245 million contract in 2022, and for the remainder of Wilson's time in Denver and several years after, they experienced below-average growth due to Wilson's underperformance and salary cap ramifications. Another example of this phenomenon is with Deshaun Watson being signed to a

$230 million deal with the Browns in 2022. Surprisingly, they experienced relatively high growth in franchise valuation the year after, but then experienced a downturn as a result of Watson not becoming a franchise player. These cases highlight that “missing” on a high-profile contract may have a negative effect on valuation, as it dims excitement for the future.

Conversely, teams that invested heavily in a successful star rarely see extensive growth. For example, the Cowboys signed Dak Prescott to a $161 million extension in 2021, but saw below average growth the following year. Meanwhile, the Chiefs signed Patrick Mahomes to a $503 million extension in 2020, but actually saw an increase in growth relative to the NFL average. This suggests that missing big on a high-profile contract may have a greater effect on a franchise valuation than rewarding success boosts it, as the damage from failing to meet expectations can negatively affect future team aspirations while success is more of an expectation.

Comparative Case Study: Chiefs vs. Raiders

To synthesize the impact of various factors influencing franchise valuation, it is critical to look at two teams that have had very different histories in the NFL. Those two teams are the Kansas City Chiefs and the Las Vegas Raiders. The Chiefs’ valuation growth has been primarily driven by the team’s legacy of success. Over the sample years of franchise valuation changes, they have had no major stadium renovations, location changes, investments in real estate, or ownership changes. However, they have not had a losing season since 2012, and have won three Super Bowls during that period. The Raiders, on the other hand, have not had as much success on the field, but off the field they have invested in a lot of revenue drivers. They moved to Las Vegas and built a new stadium in 2020 as well. They also received sponsors for their stadium and have a collaboration with Credit One Bank. Despite the Chiefs' success on the field, the Raiders have seen consistently higher valuation growth, especially in the years surrounding 2020. While the Chiefs saw moderate growth after they won more than one Super Bowl, it was not as pronounced as the Raiders surge. This comparison highlights that business-driven factors like relocation, stadium investment, and sponsorships are the most effective ways to increase valuation while on-field success only tends to result in moderate valuation growth.

Challenges

Conducting a comprehensive valuation analysis of NFL has presented several significant challenges, primarily from data accessibility online and methodological limitations.

One of the main challenges in writing this paper was a time constraint leading to some of the other major challenges. The first of which was the lack of dedicated time to more advanced coding and data modeling. Due to a lack of prior knowledge in these fields, extremely advanced metrics, graphs, and predictive models which would have been beneficial were unable to be completed. Furthermore, these technological challenges compounded by the google sheets application reduced the potential effectiveness in creating exploratory analysis.

Another challenge was finding and collecting reliable NFL financial data. Not all of the NFL's financials and other information are public, meaning disparate data sources were compiled to generate insights. Each data source had to be evaluated for veracity before inclusion in the paper, adding time and resources that could have been spent on other aspects of the paper. The wide variation in the sources with different information made it harder to define terms and ensure that the set of data that was chosen was the most accurate to the NFL. Finally, while there are strong correlations between much of the data included in this paper, direct causality was extremely difficult to prove due to data limitations.

Conclusion

This research paper, through exploratory analysis and research, has identified and ranked the most effective unshared revenue streams for increasing NFL valuations. While each stream has its pros and cons, some are clearly more effective than others, causing their impact on valuation growth to vary significantly.

For an NFL owner, the most lucrative way to increase their team's valuation is by relocating to a more populous city with a bigger market. Obviously, this is not applicable for every NFL team, such as the Giants and the Rams, who are already located in some of the biggest cities in the world and have new and modern stadiums, but for other teams, this drastic change can cause a massive spike in valuation which would be maintained for many years, as seen in the examples with the Raiders and the Rams. Furthermore, this move usually coincides with a new stadium, meaning that the team faces more potential revenue from their new venue. This method often increases fan engagement overall, and the only downside of it is that it is extremely expensive and often difficult to execute effectively.

Following relocation, developing team-owned real estate comes next, whether that is within the stadium or outside of the stadium through mixed-use developments. As teams improve their stadiums, they may opt to hold more seats or include more luxury suites that both generate large amounts of additional revenue. Also, new stadiums generally host more events outside of the NFL which can be highly efficient revenue streams. As a whole, improving or changing stadiums boosts overall fan excitement and can serve as a long-term asset. However, after enough renovations, it can be hard to increase stadium revenue. This is the main reason why additional real estate has become so popular. The revenue from the real estate is year round, goes directly to the team the real estate is anchored to, and is a long term source of revenue. The benefits of this real estate are fairly clear, as teams who have it, such as the Cowboys and the Patriots, are at the top of the league in terms of valuation.

Mid-tier revenue streams in terms of effectiveness are acquiring a stadium naming rights deal and local sponsorships. The stadium naming rights deal is easy to execute (as long as there is not an existing deal), often has no caveats, and can generally bring in over $10 million per year. It may not be the most lucrative, but the extra money can be helpful to supplement big projects. Local sponsorships, similarly, can be relatively effective in terms of revenue, but they are a little more difficult to execute than naming rights deals. Local sponsorship can boost

branding, local awareness, and increase revenue diversity, but they are not the same scale of some of the other revenue-building opportunities, and can only be extremely effective in specific scenarios.

The last, and least effective methods that an owner can do is anything related to team play or team building. In terms of team building, the best thing to do is try to acquire an extremely high caliber or popular player through the draft, free agency, or a trade. If done correctly, and only with a specific caliber player, this method can lead to a drastic spike in revenue short-term, but is not as effective due to the lack of long-term revenue. This can be seen with Tom Brady in Tampa Bay, as he ensured the team a rapid increase in valuation for the few years he was there, but the team lost that same growth after he left. When thinking about performing well during the season, it is more effective for an owner to build the best rounded team, without star stacking or acquiring the best players possible, as there is almost no correlation between number of star players and actual success. That being said, none of these factors have a significant effect on valuation, but can be tied into some of the other factors that were previously mentioned, making them relevant to mention.

Looking ahead, it is important to understand how the consumer market will change and how this will impact the NFL. By trends and recent predictions, regardless of what each team does individually, every team will continue to rise in valuation, due to the success of the NFL and the revenue sharing policy. However, NFL owners will have to get creative to try to set their team apart from the others in valuation growth. One area that has not reached a full financial maturity is the full proliferation of stadium district real-estate areas, leading into a full entertainment district. This would give the teams control over surrounding parts of the city and allow them to generate additional revenue year-round. Another idea is to launch team-controlled streaming experiences (assuming the NFL allows individual teams to own their own streaming platforms) or even digital fan experiences, as these would appeal to the increasingly digital world and be especially popular with the younger consumers. Lastly, NFL owners should consider expanding internationally to try and gain a larger fanbase, while working within restrictions placed upon them. This would be like having a secondary home stadium, in a sense, in another country to try to gain the people there as support for the franchise. By doing this expansion, team owners can acquire revenue from two different home stadiums and markets, which has the potential to double NFL valuations in a short period of time. The recently added international NFL game schedule is evidence of them trying this already.

These dynamics and methodologies reflect a broader evolution in sports across the world overall, but with the NFL's unique levels of popularity, it will continue to be a dominant leader in sports business for years to come. In today's NFL, success belongs not to the teams that win most, but to the franchises that operate as diversified, risk-savvy businesses beyond the field.

Works Cited

"2024 NFL All-Pros." Pro, www.pro-football-reference.com/years/2024/allpro.htm. Accessed 24 July 2025.

"2024 NFL Franchise Values: See Where All 32 Teams Rank." NBC 5 Dallas-Fort Worth, NBC 5 Dallas-Fort Worth, 5 Sept. 2024, www.nbcdfw.com/news/sports/nfl/nfl-franchise-values-2024-full-list/3636564/.

Badenhausen, Kurt. "How NFL Teams and Owners Make Their Money." Yahoo! Sports, Yahoo!, sports.yahoo.com/nfl-teams-owners-money-110000452.html?guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAN7N5FqgSmcZc4vy-raMo2XNv9P9gubXlXNPFgRyg88Vcayca6FGAqCb2os3F7rMVS3dwoQ0lcfQuE1H0LWapTcVC9Oi-oUhkjy1fLCzvo2Wf4iut31-vUQ_6OzOnl5jghf6RziFj-St_rRU2sjAYWqEYGOuqZT9JfKyXaxIy_j7. Accessed 24 July 2025.

Badenhausen, Kurt. "NFL Playoffs Generate No Financial Windfall for Individual Teams." Sportico.Com, Sportico.com, 14 Jan. 2022, www.sportico.com/leagues/football/2022/nfl-playoffs-generate-no-financial-windfall-1234658372/

Beach, Michael. "The NFL Is Back for More Cash." Cross Screen Media, 6 Sept. 2024, crossscreen.media/state-of-the-screens/the-nfl-is-back-for-more-cash/.

Bisson, James. "NFL Naming Rights Sponsorships: Which Teams Earn the Most Money?" Sportsbook Review, 10 July 2025, www.sportsbookreview.com/news/nfl-sponsorships/.

Bisson, James. "NFL Owner's Suite Rankings: Cost Breakdowns for All 32 Teams." Covers.Com, 17 July 2025, www.covers.com/nfl/most-expensive-owners-suites.

Cridlin, Jay. "How Tom Brady's Return Impacts Tampa Bay's Economy, from Tickets to Tourists." Tampa Bay Times, www.tampabay.com/news/business/2022/03/23/how-tom-bradys-return-impacts-tampa-bays-economy-from-tickets-to-tourists/. Accessed 25 July 2025.

"Dallas Cowboys on the Forbes NFL Team Valuations List." Forbes, Forbes Magazine, www.forbes.com/teams/dallas-cowboys/. Accessed 24 July 2025.

Dent, Mark. "Are Superstar Athletes Good for the Economy?" The Hustle, HubSpot, 9 Feb. 2024, thehustle.co/are-superstar-athletes-good-for-the-economy.

Eckstein, Jakob. "How the NFL Makes Money." Investopedia, Investopedia, www.investopedia.com/articles/personal-finance/062515/how-nfl-makes-money.asp

ESPN, ESPN Internet Ventures, www.espn.com/nfl/attendance. Accessed 24 July 2025.

"Field Position: Does an NFL Stadium Boost Your Home Value?" Trulia, 21 May 2018, www.trulia.com/research/nfl-stadium/.

FinanceBuzz. "Financebuzz Tracked Prices for Tickets, Concessions, and Parking for Every National Football League (NFL) Team from 2013 to 2023, to See Which Teams Raised Prices the Most and Least.: Financebuzz." FinanceBuzz Tracked Prices for Tickets, Concessions, and Parking for Every National Football League (NFL) Team from 2013 to 2023, to See Which Teams Raised Prices the Most and Least. | FinanceBuzz, 5 Sept. 2024, www.linkedin.com/posts/financebuzz_nfl-inflation-activity-7237442086503870466-2Ybn

"Forbes List of the Most Valuable NFL Teams." Wikipedia, Wikimedia Foundation, 9 July 2025, en.wikipedia.org/wiki/Forbes_list_of_the_most_valuable_NFL_teams.

Harrison, Steve. "NFL Teams Are Building State-of-the-Art Practice Palaces. Is 'Panthers Park' Coming?" The Charlotte Observer, www.charlotteobserver.com/sports/nfl/carolina-panthers/article213182114.html. Accessed 26 July 2025.

Johnson, Matt. "NFL Owners Reportedly Not Happy over Rising Costs amid Record-Setting Revenue." Yahoo! Sports, Yahoo!, 25 May 2025, sports.yahoo.com/article/nfl-owners-reportedly-not-happy-150000403.html.

Johnson, Matt. "NFL Revenue: Examining Where the National Football League's Money Comes from, Future of NFL TV Revenue." Sportsnaut, 26 May 2025, sportsnaut.com/nfl/where-does-nfl-revenue-come-from/.

"List of Current NFL Stadiums." Wikipedia, Wikimedia Foundation, 4 July 2025, en.wikipedia.org/wiki/List_of_current_NFL_stadiums.

"List of Kansas City Chiefs Seasons." Wikipedia, Wikimedia Foundation, 17 July 2025, en.wikipedia.org/wiki/List_of_Kansas_City_Chiefs_seasons.

Laughlin, Alex. "The Cost (and Dangers) of Buying a Quarterback | Sports Analytics Group at Berkeley." Sports Analytics Group Berkeley, sportsanalytics.studentorg.berkeley.edu/articles/quarterback-cost.html. Accessed 24 July 2025

Liu, Yazhe. "THE STAR EFFECT ON SPORT TEAM REVENUE: EVIDENCE FROM THE NATIONAL BASKETBALL ASSOCIATION." Washington State University, Aug. 2012, www.google.com/url?sa=i&url=https%3A%2F%2Frex.libraries.wsu.edu%2Fview%2FpdfCoverPage%3FinstCode%3D01ALLIANCE_WSU%26filePid%3D13338212630001842%26download%3Dtrue&psig=AOvVaw1Zol_sueZeq8lJwtSHzHBM&ust=1750354903778000&source=images&cd=vfe&opi=89978449&ved=0CAYQrpoMahcKEwjw0eejwvuNAxUAAAAAHQAAAAAQBA .

Loignon, Andy. "Why Successful Leaders Are Wary of the 'Too Much Talent Effect.'" Big Think, 10 Jan. 2024, bigthink.com/business/why-successful-leaders-are-wary-of-the-too-much-talent-effect/.

Klawans, Justin. "The Economics of Taxpayer-Subsidized Stadiums." The Week, The Week, 23 May 2024, theweek.com/sports/taxpayer-subsidized-stadiums.

Moore, Camille. "Most Valuable NFL Teams: A Comprehensive Ranking for 2023." Worthly, 27 Nov. 2023, worthly.com/value/most-valuable-nfl-teams-2/.

Mueller, Kristin. "When Sports Facilities Anchor Mixed-Use Developments, Everyone Wins." Forbes, Forbes Magazine, 25 Oct. 2024, www.forbes.com/sites/kristinmueller/2024/10/24/when-sports-facilities-anchor-mixed-use-developments-everyone-wins/.

"NFL Franchise Moves and Mergers." Wikipedia, Wikimedia Foundation, 23 July 2025, en.wikipedia.org/wiki/NFL_franchise_moves_and_mergers.

"NFL Game Cost Inflation over the Last 10 Years." Voronoi by Visual Capitalist, www.voronoiapp.com/sports/NFL-Game-Cost-Inflation-Over-the-Last-10-Years--2708.

Accessed 24 July 2025.
"NFL Sponsorship Trends: A Deep Dive into Revenue Distribution, Retail Demand, and Super Bowl Activations." SponsorUnited, www.sponsorunited.com/insights/nfl-sponsorship-trends. Accessed 25 July 2025.
"NFL Team Win Trends - All Games, since 2013." NFL Team Win Trends - All Games, www.teamrankings.com/nfl/trends/win_trends/?range=yearly_since_2013. Accessed 24 July 2025.
Ross, Stephen, and Benjamin Woodworth. The Basic Revenue Sharing Issue, pennstatelaw.psu.edu/_file/Sports%20Law%20Policy%20and%20Research%20Institute/More_Like_the_United_Nations.pdf. Accessed 25 July 2025.
"Saquon Barkley TOPS NFLPA's Year-End Top 50 Player Sales List." NFL Players Association, nflpa.com/partners/posts/things-to-know-about-the-2024-year-end-nflpa-top-50-player-sales-list. Accessed 25 July 2025.
Shrikhande, Amol. "NFL Metro Areas by Population." ComposeMD, 17 Sept. 2024, composemd.com/2023/09/18/nfl-metro-areas-by-population/.
Talkington, Erin. "Sports Venue-Anchored Real Estate Development – Overview by League - RCLCO Real Estate Consulting." RCLCO Real Estate Consulting - The Best Minds in Real Estate, 1 Oct. 2024, www.rclco.com/publication/sports-venue-anchored-real-estate-development-overview-by-league/
Tamir, Alon. "The Value of Winning: Sports Analytics Group at Berkeley." The Value of Winning Sports Analytics Group at Berkeley, sportsanalytics.studentorg.berkeley.edu/articles/value-winning.html. Accessed 24 July 2025.
"'The Stadium Is Secondary': How US Sports Teams Became Real-Estate Speculators." The Guardian, Guardian News and Media, 6 Dec. 2023, www.theguardian.com/sport/2023/dec/06/mixed-use-development-sports-teams-owners-stadiums
The Top Ranked NFL Quarterbacks by 2010s Season, ainsworthsports.com/football_player_rankings_by_season_nfl_qb_2010s.htm. Accessed 24 July 2025
The Top Ranked NFL Quarterbacks by 2020s Season, ainsworthsports.com/football_player_rankings_by_season_nfl_qb_2020s.htm. Accessed 24 July 2025.
Thomas, Will. "Is Too Much Talent Damaging Your Team's Performance?" LinkedIn, 8 Mar. 2018 www.linkedin.com/pulse/too-much-talent-damaging-your-teams-performance-will-thomas-phd.
"USF Scholarship: A Digital Repository @ Gleeson Library: Geschke Center." Site, repository.usfca.edu/. Accessed 24 July 2025.
"What Are Some of the Factors That Determine How Much an NFL Franchise Is Worth? - Appraisal Economics." Appraisal Economics - Valuation Services, 18 Nov. 2020, www.appraisaleconomics.com/factors-determine-much-nfl-franchise-worth/.
White, Shannon. "Even Elite Franchise Quarterbacks Need a Strong Supporting Cast." Behind the Steel Curtain, Behind the Steel Curtain, 23 Dec. 2022, www.behindthesteelcurtain.com/2022/12/23/23520197/even-elite-franchise-quarterbacks-need-a-strong-supporting-cast-steelers-kenny-pickett-nfl-news.

Healing with Machines: Advancements, Challenges, and Ethical Considerations of AI in Healthcare By Aadya Bansal

Abstract

As the presence of artificial intelligence (AI) and automation increases in healthcare, many questions are raised about the balance between technological advancements in the fields and the quality of patient care. This paper explores the extent to which healthcare should be automated so that it improves the quality of care and while minimizing harm to the patient. AI powered tools bring many advantages to the healthcare field such as a more accurate and faster diagnosis and treatment recommendations, higher surgical precision, and the ability to reduce administrative burden by managing repetitive tasks. However, there are many concerns regarding the ethical implications of such technology, a lack of patient trust and transparency and the risk of skill degradation among healthcare professionals. In this paper we evaluate both the benefits and risks of automation to conclude how AI can best human healthcare professionals instead of completely replacing them in patient care. By analyzing AI's existing use in medical fields, future trends, and ethical dilemmas, this paper emphasizes the need for a balanced approach for implementing AI in healthcare in which AI's strengths are leveraged with proper human oversight.

1. Introduction

The integration of AI in healthcare has been transformative for many medical practices. AI systems show promise in helping increase the accuracy of diagnoses, treatment plans and helping streamline repetitive administrative tasks. Sophisticated AI systems show promise in enhancing the precision of diagnoses, treatment plans and streamlining many administrative tasks (Grenier et al., 2022). Healthcare, independent of AI use, is largely reliant on the expertise of professionals, their empathy and the speed of real-time decision-making. However, rising patient loads, medical errors, and inefficiencies in healthcare delivery have necessitated technological intervention (David Cutler et al., 2020). AI has shown its ability to process huge and complex datasets, identify patterns, and make decisions at unprecedented speeds (Liang & Hu, 2015). When applied to medicine, AI presents many opportunities but also risks regarding its lack of transparency and confusion about accountability related to medical decisions (Ueda et al., 2024).

This paper will examine the current applications of automation and AI technologies in healthcare, like AI-assisted diagnostics, robotic surgery, and remote patient monitoring. Next, it will explore the future uses of AI, upcoming technologies and their potential to become more autonomous in a healthcare setting. This will then lead to a discussion addressing the issues relating to different ethical, legal, accountability and trust. Finally, the paper will evaluate how much automation can be implemented responsibly without compromising patient security and care.

2. Current Automation in Healthcare

Automation in healthcare has started to become sophisticated, especially with a variety of AI-driven devices and systems. These technologies are prominently being used in data-intensive fields such as dermatology, oncology, radiology, pathology, and ophthalmology. Currently, their primary function has been to assist doctors by suggesting a diagnosis, detecting findings and recommending treatments. By processing vast amounts of data quickly and accurately, these systems help in human decision-making, improving the care delivered. They also support doctors in other clinical tasks such as triage, by assisting with prioritizing time-sensitive cases and monitoring a patient's trajectory. As of now, most systems are assistive and present in acute care settings in high-income countries. Over time, though, they are expected to become autonomous by going beyond giving just a recommendation, to operating closed-loop machines like ventilators and insulin pumps. Immediate effects on medical professionals using AI include a significant workload to review and confirm AI output and a reduction in time spent per patient. The long-term effects are predicted to be situational awareness and skill degradation, which are well-documented effects of automation in other industries

AI in Clinical Decision Support

One of the biggest applications of AI is helping clinicians make a faster and a more accurate diagnosis and provide treatment plans. During times of a high patient workload, they can reduce the workload on clinicians by helping analyze a patient's medical history and any other related data.

For example, in radiology, AI systems can assist doctors by acting as a second pair of eyes in analyzing medical images such as X-rays, CT scans, and MRIs, to catch any abnormalities that may have been missed, such as lung nodules, cardiovascular calcifications, and degenerative spinal changes. Siemens Healthineers recently developed a deep learning model called AI-RAD Companion that can be used by radiologists. By having a quick second reading, radiologists gain confidence in their diagnosis, shorten turnaround time, provide better patient outcomes, and eventually reduce costs (Grenier et al., 2022). Similarly, in ophthalmology, Gulshan et al. (2016) describes how Google's AI system was developed and validated for the detection of diabetic conditions such as diabetic retinopathy and age-related macular degeneration. Deep learning was also used in this model so that the algorithm could analyze retinal images and identify early signs of the disease, resulting in quick intervention. In dermatology, applications (apps) are becoming more prominent for diagnosing different conditions. Apps like VisualDx provide patients with easy to understand personalized answers, while apps like Miiskin help patients keep track of their skin and moles and store the data so that it can be shown in-person later.

Robotic Surgery

Robotic systems represent another significant area of automation in the healthcare industry. Most robotics is used mostly for surgery since it provides high-quality treatment by

bringing higher accuracy and precision compared to human surgeons. If the current trend of robotics being used continues in healthcare, the global medical robotics market systems are projected to grow from $16.0 billion (2024) to $33.8 billion in 2029, growing at a CAGR rate of 16.9% (MarketsandMarkets, 2024).

In Verma et al.'s (2018) review of IoT and robotics, they lay out many recent advancements in the field. Robotics has been deeply intertwined in healthcare, prominently through surgery, since the 1990s. Today, surgical robots comprise more than 64% of all medical robotic systems. Currently, most robots assist surgeons and have proved that robotic surgery can reduce surgical trauma and loss of healthy tissue, leading to a quicker recovery of the patient. They achieve this by working with micro-precise movements and without tremors. There are two main systems in robotics that assist in surgery, called Robotic assisted surgery (RAS) and computer assisted surgery (CAS). RAS is a motorized system controlling the motion of intervention. It consists of a robot which can be manipulated by the surgeon. RAS provides human dexterity and ergonomics while reducing tremors and resisting human errors, making it more reliable. CAS is a computer interface that assists the doctor in controlling the intervention site manually. It is generally used for detection tools, motion planning, identification, cancer cells etc.

The da Vinci Surgical System is one of the most well-known RAS systems. It consists of a console where the surgeon sits and controls the robot's arms, which hold surgical instruments. The system also provides a high-quality, 3-D view of the surgical site and the robot's arms can make micro-precise movements, which are often more stable than a human hand and reduce the risk of errors during surgery. Another RAS system is the Mako Robotic-Arm Assisted Surgery System, which is used in orthopedic procedures. This system also provides the surgeon with a 3D model of the patient's anatomy, like the da Vinci system, which allows them to plan the surgery better and conduct it with more precision. A more precise surgery means better patient results, such as reduced pain and a quicker and faster recovery. Robotic systems are also making advances in neurosurgery. The Robotic Stereotactic Assistance (ROSA) system is designed to assist neurosurgeons in performing brain surgeries with minimal invasiveness. It helps in the precise placement of instruments during procedures such as biopsies, electrode implantation, and other neurosurgical interventions which are important for such a delicate organ.

Internet of Things (IoT) in Healthcare

The Internet of Things (IoT) is a series of connected devices and key players in driving healthcare automation by helping patients collect and monitor data. Patients can do this by using new devices and services being created by IoT startups and many tech giants like IBM, Google, Apple, and Samsung. These devices are gaining traction in this area by commercializing devices and services that assist in user health by helping patients get their health data on a daily basis.

One of the most prominent applications of IoT in healthcare is remote patient monitoring. IoT enabled devices such as smartwatches, wearable electrocardiogram (ECG) monitors and continuous glucose monitors, allow doctors to monitor their patient's data without needing to call

them in, saving both time and money (Park et al., 2020). For example, in 2017, Apple installed an FDA-approved deep learning algorithm in their smartwatch that detects atrial fibrillation (a common type of arrhythmia). Using photoplethysmography and accelerometer sensors, it learns the user's average heart rate at rest and at activity, and sends a warning signal if the heart rate becomes abnormal (Nguyen et al., 2018). Another IoT application is medication management. Smart pill dispensers, for example, can be programmed to release the correct dosage on time, reducing the number of missed or incorrect doses. Since economic advantages are the main driving factors of IoTs in healthcare, their implementation results in better outcomes, disease and drug management with cheaper cost of treatments.

3. Future of AI in Healthcare

Detection

AI algorithms like recurrent CNNs– a type of CNN that uses recurrent connections within its layers which allow the network's units to evolve over time, making them suitable for tasks that require contextual information, like object recognition (Liang, Hu, 2015)– can be used to analyze diseases like coronary atherosclerotic plaques which usually require an invasive and demanding procedure like (Zachariadis, Leligou, 2024). A newer, less invasive technique of testing for abnormalities in blood flow and diagnosing cardiovascular diseases is coronary computed tomography angiography (CCTA). CCTA is non-invasive and creates 3D images of the heart (Ramjattan et al., 2024) that doctors can use with a possibility of overestimating the degree of narrowing of the artery lumen (Zachariadis, Leligou, 2024). To improve its diagnostic accuracy, Federate Learning (FL) has been applied in this context. FL helps institutions collaborate to train AI models on a diverse range of patient data without any direct sharing of sensitive patient data which significantly affects CCTA's diagnostic performance (Di Cosmo et al., 2023). By incorporating data from multiple sources, FL can reduce biases in the model that stem from training a model on a single dataset, which leads to models that are better adapted for different demographics.

Similarly, in musculoskeletal imaging, AI matches or surpasses human expertise in detecting fractures and degenerative conditions. The Sytin model (VGG-16) was trained on hip X-rays and achieved a sensitivity of 95% and an overall accuracy of 92.8% in diagnosing hip osteoarthritis from X-rays, making it comparable to physicians who have atleast 10 years of experience (Zachariadis, Leligou, 2024). AI's capabilities can go beyond osteoarthritis detection. For example, faster RCNN models have shown an accuracy of 88% in detecting interochanteric than orthopedic specialists, who had an accuracy of 84% (Liu et al., 2022). Additionally, ResNet18 has shown an accuracy of 98.1% in identifying distal radius fractures, (Tobler et al., 2021), further proving AI's ability to assist radiologists in high-precision musculoskeletal assessments. However, while AI excels in detecting fractures, it performs moderately in complex classification tasks such as evaluating joint involvement and fragmented displacement with an accuracy ranging from 58.9% to 85/1% in the task (Yoon et al., 2020).

Deep Learning Applications in Surgery

Deep learning (DL) is another tool for enhancing surgical practices by helping increase precision, introducing robot-assisted interventions and automated procedural recognition. An application of DL is automated phase recognition in which models like EndoNet identify different surgical stages to provide real-time feedback that helps improve surgical safety and give skill assessments to residents (Yang et al., 2024). Region-based CNNs have the capability to localize surgical instruments while deep reinforcement learning can refine robotic techniques through continuous feedback, leading to improved precision and efficiency (Morris et al., 2022).

DL can be applied to various surgical specialties: in orthopedics, DL models have a high accuracy in diagnosing fractures and detecting orthopedic implants from radiographs (Bozzo et al., 2024); in neurosurgery, it can identify anatomical structures and track the surgical phases to minimize the risk of a complication during the surgery, thereby improving patient safety (Morris et al., 2022). Similarly, in plastic surgery, DL can assess and evaluate multiple postoperative results. Applications like DL4Burn analyze visual data to help in burn surgery, providing more effective treatment strategies (Morris et al., 2024).

IoT

The integration of IoT into healthcare also has a huge scope for improving patient outcomes, especially through real-time monitoring. By interconnecting different devices and systems, IoT can help establish a seamless connection between medical tools, wearables, and healthcare providers (Zaman et al., 2024). IoT in healthcare would operate across multiple layers which would manage data collection, transmission and application. In the perception layer, there would be technologies like RFID tags, medical sensors, and GPS devices to collect real-time health data. These tools would be responsible for converting physical phenomena into digital signals that would go to the patient and potentially the devices of their healthcare providers, who would then be able to analyze the metrics (Zaman et al., 2024). For example, devices with biometric sensors that can track heart rate, blood pressure, and oxygen levels allow for non-stop monitoring without hospital visits.

The network layer would primarily be used to manage the transmission of the collected data from the perception later using Bluetooth and Wi-Fi. A strong network layer is necessary in cases of emergency situations where real-time data can help in rapid decision-making and intervention. However, data storage and security are two main concerns in the area. Even though cloud-based storage is flexible and scalable, there are significant concerns raised regarding patient data privacy during potential security breaches. To address these concerns, technologies like edge computing and blockchain technology are being explored to create a decentralized approach. Edge computing would allow real-time data processing closer to the source, which also reduced latency and bandwidth usage. On the other hand, blockchain is more focused on patient autonomy and reducing risks related to centralized data breaches (Nguyen et al., 2024). Lastly, the application layer would use AI to interpret and analyze the health data. AI algorithms

can analyze electronic health records, data and other patient information to assist healthcare professionals in making a diagnosis (Kelly et al., 2020).

Apart from diagnosis, IoT technologies could also play a role in increasing healthcare accessibility in underserved areas. By using monitoring and telemedicine, patients would be able to receive medical consultations and follow-ups without physically having to visit a hospital or clinic (Pullimamidi et al., 2022). Wearable health trackers and mobile applications would promote a disease prevention approach by allowing individuals to monitor their own health and be alert in case of any abnormal changes (Kelly et al., 2020). The integration of IoT can help healthcare shift towards a more efficient and accessible patient care model. However, strong data security measures are needed to ensure safe use and trust in the technology.

4. Advantages and Limitations

In the healthcare system, the integration of technology has both significant advantages and limitations.

Automation can significantly reduce the time needed by clinicians and staff to do time-consuming and repetitive administrative tasks, which improves both efficiency and the quality of patient care. In the U.S., about one quarter of the total healthcare expenditure is used for administration, such as getting authorizations, eligibility inquiries and managing billing processes (David Cutler et al., 2020). Automation through technologies like RPA can streamline these processes that traditionally require a large amount of manual labor. For example, in a case study from National Taiwan University Hospital, the time to process medical claims was reduced by 380 minutes through the implementation of RPA, bringing up the cycle efficiency to 95.54% from 69.07%. Therefore, this integration would help healthcare professionals give more attention to achieving faster turnaround times and patient centered care instead of on non-value steps like filtering and manual data entering (Huang et al., 2024).

The integration of AI also shows promise in being able to enhance operational efficiency. Tools like clinical decision support systems help in improving diagnostic accuracy and optimizing workflows. For example, since AI-driven tools can rapidly analyze a patient's data and provide suggestions on a diagnosis and treatment plans, clinicians can make informed decisions faster, which in turn reduces wait times. Moreover, these predictive tools can help identify a high-risk population early by analyzing risk factors, genetics, lifestyle choices, environmental exposures, and existing health conditions (Khalifa, Albadawy, 2024), giving them the opportunity for early proactive interventions which can also reduce hospital admissions and a sudden increase in workload (Kasula, 2023).

Increased diagnostic accuracy is another benefit of automation technologies. Because they can perform an analysis of vast amounts of patient data, AI algorithms are excellent at identifying subtle patterns that might be missed by clinicians, especially when they're under a heavy workload. For example, if there is an AI continuously monitoring vitals, it can flag small abnormalities that could indicate a patient's deteriorating condition (Yelne et al., 2023).

Therefore, using these technologies to minimize human errors would lead to timely and accurate diagnosis, and ultimately, take patient care to another level (Venigandla, 2022).

While automation brings clear advantages to healthcare, these benefits come with certain challenges. The biggest concerns are caused by the lack of human interaction that the increased implementation of AI technologies would cause and the "black box" nature of AI algorithms. These factors then raise questions about transparency, accountability and trust, especially when critical patient decisions are based on opaque, automated inhuman algorithms.

The reduction in human interaction can negatively affect the patient's experience. As technology increasingly handles tasks that take place through interaction with doctors, such as making a diagnosis, patient monitoring and decision-making, the human elements – empathy and real-time communication – diminish. This can lead to a feeling of isolation or even distrust among patients, who may need reassurance and interaction rather than blind trust in an algorithm. Moreover, algorithms are more likely to overlook context-driven factors in patients that only humans can detect (Abad et al., 2010). A study by Digby et al. (2024) indicates the importance of human interaction in patient care, especially when it comes to those who are in isolation. Many negative patient experiences stem from a lack of interaction during their treatment which affects their physical and mental health and may lead to increased levels of depression, anxiety and fear. So, as the duration of their isolation lengthens, the negative emotional states also intensify, which is proved by a worsening Hospital Anxiety and Depression Scale over time. Even though some individuals prefer increased privacy, the overall impact is overwhelmingly detrimental, highlighting the need for a way to address the mental health of patients when they have fewer interactions with their healthcare providers (Abad et al., 2010).

The "black box" problem with deep learning AI technologies may arise because it is impossible to know how the system came to its conclusion, which hinders their acceptance in clinical settings. The lack of transparency in their decision-making leads to skepticism and diminished trust in both healthcare providers and patients despite the statistics regarding AI's high performance in diagnostic accuracy. Medical practitioners require transparency on matters like clinical performance, safety and risks because they are liable when using AI tools on their patients. On the other hand, patients want transparency because they have a right to know whether the predictions made by the AI software are safe for them (Fehr et al., 2024). In a survey conducted by Tyson et al. (2023), six in ten U.S. adults stated that they would feel uncomfortable if their doctors relied on AI for diagnosing them and recommending treatments. There are a variety of factors that lead to this distrust. There is a major concern about AI's impact on the relationship between the patient and their healthcare provider: 57% say that the relationship will only get worse. The security of health records is another concern, with 37% believing that AI would increase the security risk.

This lack of clarity among the public makes it nearly impossible to understand how these models arrived at their conclusion, which affects the safety of the patient. Moreover, the earlier use of any ML-based model may disproportionately affect certain patient populations, since the models could be trained on datasets that are inherently biased. A study by Obermeyer et al.

(2019) highlights how racial bias was identified in an algorithm used to predict a patient's healthcare needs. The algorithm used healthcare costs as a proxy for medical needs and consistently overestimated the severity of the illness in white patients compared to black patients. It only recommended enhanced care for 17.7% of Black patients compared to the 46.5% of white patients, even though both were at the same risk level. This bias originates because the algorithm uses healthcare costs as a proxy for health needs, but Black patients have historically spent less on medical expenses than white patients for their health needs. This problem arises from systemic issues like different socioeconomic backgrounds, difficulty in accessing healthcare, and direct or indirect racism in the healthcare system, and it is hard to address this discrepancy without insight into the model's internal workings (Jhaveri et al., 2022).

5. Impacts

Building on the advancements of AI, it is also crucial to address the broader impact of the automation of healthcare delivery. These innovations promise early disease detection, lowering clinical loads, and help in times of resource shortages to help get better patient outcomes. These innovations do promise better patient outcomes through early disease detection, lower clinical loads and addressing other resource shortages in the healthcare system. However, it will likely face major challenges related to patient trust in technology, data privacy and a lack of transparency. The following sections examine ways AI and automation will change healthcare on a broader level.

Enabling Faster Diagnoses

Automation and AI-driven systems can enhance patient care quality by speeding up making a diagnosis and a treatment recommendation, leading to quicker and more informed decisions. For example, they can help analyze medical images, X-rays, MRI and other scans in real-time, which usually take 24-48 hours. They can also act as a second pair of eyes to help clinicians catch any missed details in time constrained environments (Grenier et al., 2022). This early detection and treatment can enable prompt interventions therefore contributing to improved long-term health outcomes because the conditions are managed before they have a chance to progress to severe stages.

Automation Creates More Personalized Healthcare

Another way that automation technologies can greatly improve patient care is by helping create more personalized care plans by analyzing a patient's health-related data such as their health records, genetic data and data from wearable devices. For example, the Apple Watch has FDA approved deep learning algorithms to detect atrial fibrillation, monitor heart health and alert the user to any abnormalities it detects(Nguyen et al., 2018). Similarly, Continous Glucose Monitors (CGMs) help diabetic patients maintain glucose levels by giving them their blood sugar trends (Park et al., 2020). Therefore, personalized healthcare can help improve patient

outcomes by addressing their individual needs, but strong security measures are needed to make sure that the patient's data is kept safe.

At-Home Care Expands Access to Treatment

Automation and AI-powered devices are also making at-home care more available to expand the range of access to treatments for patients. Devices like wearable ECG monitors, CGMs, and smart pill dispensers allow healthcare providers to monitor their patients' health remotely, which reduces the need for frequent hospital visits. This makes patient care more convenient and continuous, which is helpful for patients with chronic conditions that need constant monitoring. But there are challenges in scaling at-home care, such as the digital divide due to which patients in low-income or in areas that lack access to such devices or a reliable connection that can connect them to their doctors; or a patient's lack of knowledge about the proper usage of specific devices can be an obstacle in them getting proper care because these devices rely on proper usage and care which can vary widely (Park et al., 2020).

Automation Reshapes Healthcare Roles

The introduction of healthcare will have a huge impact on healthcare professionals by changing their roles and requirements. While automation can handle mundane and repetitive tasks like data entry, imaging analysis and scheduling, the clinicians will be able to spend more time working on complex cases and patient interaction. However, this transformation will also raise concerns regarding workforce displacement. Since some redundant tasks will be taken over, there will be the need to reskill or upskill the workforce that took care of them before. On the other hand, the effect of a shortage of healthcare professionals in certain fields like radiology and pathology can be reduced by supporting clinicians in diagnosing and monitoring patients. For example, AI tools can assist overburdened radiologists by acting as a second pair of eyes to ensure that no details are missed. Similarly, in areas that lack specialists, these tools can assist general physicians or lower-level healthcare workers in making a diagnosis.

6. Ethics

As AI technologies are increasingly integrated into the healthcare system, ethical questions surrounding its liability, lack of transparency, and regulation rise. This section explores the ethical dilemmas related to AI and the current efforts to establish guidelines for its responsible implementation.

One major issue is determining liability for when AI systems cause harm while operating is an ethical and legal gray area. The dominant view is that the developers and manufacturers of the product should be held responsible in this case. This is also consistent with product liability laws such as the Restatement (Third) of Torts: Product Liability, which holds the manufacturers accountable for any harm caused by defects in their products such as design flaws and inadequate instructions (H. Silverglate, 2001). If the AI system makes an incorrect diagnosis due to flawed training data or programming, the developers may be held responsible as well.

According to a poll conducted by the AI Policy Institute, 73% of voters believe that AI companies should be held responsible for the harms caused by their technologies (Colson, 2023), highlighting the public support for this approach.

Another approach to the problem is to hold healthcare providers using AI systems equally responsible. According to malpractice laws, clinicians should exercise their judgement and evaluate the AI recommendations before applying them to the patient. The Health Information Technology for Economic and Clinical Health (HITECH) Act emphasizes the shared accountability of healthcare providers in managing and applying any technology. But many critics of this approach argue that not all clinicians may have the necessary knowledge to fully understand the reasoning behind an AI model and AI's "black box" characteristics would only make it harder for them to do that, and therefore cannot.

This issue of liability becomes more pressing when an AI tools recommendation or results conflict with those of a human professional. The view that the medical professional's decision should be final stems from the common belief that they are more equipped to take into account a patient's contextual factors like their preferences, mental state and culture. This also aligns with the idea of a patient-centered healthcare system in which human understanding and empathy play a huge role. In a 2023 survey by Tyson et al. 60% of U.S. adults said that they would feel somewhat or very uncomfortable about their healthcare provider relying on AI for diagnosing and recommending treatments, underscoring a strong public preference for human oversight in medical decisions.

On the other hand, since AI systems base their recommendations on vast datasets and evidence that clinicians might not have the time or capability to process, they are more capable of having a higher accuracy and identifying early signs of a disease. For example, AI systems in radiology have outperformed human clinicians in identifying early stages of cancers (Grenier et al., 2022). In cases like this, a clinician's decision might be detrimental to a patient. Therefore, it becomes important to balance both judgements. And to achieve a proper balance, there is a need for structured protocols and rules like the ones in the UK's National Health Service. The NHS recommends the implementation of an environment in which AI is a tool for decision support, but the clinician has the final accountability.

7. Conclusion

The integration of AI into healthcare is one of the most significant technological shifts in medicine to date. While AI has already enhanced diagnostics, treatment planning and administrative tasks, there is constant debate over the extent to which this should be continued. Therefore, the challenge isn't deciding whether AI and automation technologies should be implemented but how they can be used best to support human expertise while upholding ethical integrity.

One of the concerns is the risk of automation eroding human aspects of patient care such as human empathy and clinical judgement, which take contextual factors into account. An overreliance on AI could lead to skill degradation among healthcare professionals, which makes

it crucial to have a balance between technological efficiency and the current medical institution. Additionally, the effects of the integration of such technologies on employment in healthcare, regulatory policies and data security are all areas of concern.

Therefore, the success of AI and automation technologies in healthcare rests on policymakers, medical professionals and AI developers, who need to continuously work on establishing frameworks to mitigate AI's risks to patients and enhance its benefits as technology evolves. So, rather than replacing human decision-making, which has been imperative in healthcare, AI should function as an assistive tool for them to expand access to high-quality care. The future of healthcare automation will be dependent not just on technological advancement but on the ethical decisions made to ensure patient safety and trust.

Works Cited

Bozzo, A., Tsui, J. M. G., Bhatnagar, S., & Forsberg, J. (2024). Deep Learning and Multimodal Artificial Intelligence in Orthopaedic Surgery. *The Journal of the American Academy of Orthopaedic Surgeons*, *32*(11), e523–e532. https://doi.org/10.5435/JAAOS-D-23-00831

Colson, D. (2023, September 13). *Poll shows voters oppose open sourcing AI models, support regulatory representation on boards, and say AI risks outweigh benefits*. Artificial Intelligence Policy Institute. https://theaipi.org/poll-shows-voters-oppose-open-sourcing-ai-models-support-regulatory-representation-on-boards-and-say-ai-risks-outweigh-benefits/

Cutler, D. M. (2020). Reducing administrative costs in US health care. *The Hamilton Project Policy Proposal*, *9*, 3-25.

Di Cosmo, M., Migliorelli, G., Francioni, M., Muçaj, A., Maolo, A., Aprile, A., ... & Moccia, S. (2023, September). A Federated Learning Framework for Stenosis Detection. In *International Conference on Image Analysis and Processing* (pp. 211-222). Cham: Springer Nature Switzerland.

Grenier, P. A., Brun, A. L., & Mellot, F. (2022). The potential role of artificial intelligence in lung cancer screening using low-dose computed tomography. *Diagnostics*, *12*(10), 2435.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *jama*, *316*(22), 2402-2410.

Kelly, J. T., Campbell, K. L., Gong, E., & Scuffham, P. (2020). The Internet of Things: Impact and implications for health care delivery. *Journal of medical Internet research*, *22*(11), e20135.

Liu, P., Lu, L., Chen, Y., Huo, T., Xue, M., Wang, H., ... & Ye, Z. (2022). Artificial intelligence to detect the femoral intertrochanteric fracture: the arrival of the intelligent-medicine era. *Frontiers in Bioengineering and Biotechnology*, *10*, 927926.

MarketsandMarkets. (2024). *Medical robotic systems market - Global forecast to 2029.* Retrieved from https://www.marketsandmarkets.com/Market-Reports/medical-robotic-systems-market-2916860.html

Ming Liang and Xiaolin Hu, "Recurrent convolutional neural network for object recognition," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 3367-3375, doi: 10.1109/CVPR.2015.7298958.

Morris MX, Rajesh A, Asaad M, Hassan A, Saadoun R, Butler CE. Deep Learning Applications in Surgery: Current Uses and Future Directions. The American SurgeonTM. 2022;89(1):36-42.

Nguyen, T., Nguyen, H., & Gia, T. N. (2024). Exploring the integration of edge computing and blockchain IoT: Principles, architectures, security, and applications. *Journal of Network and Computer Applications*, 103884.

Park, C. W., Seo, S. W., Kang, N., Ko, B., Choi, B. W., Park, C. M., ... & Yoon, H. J. (2020).

Artificial intelligence in health care: current applications and issues. *Journal of Korean medical science*, *35*(42).

Pulimamidi, R., & UI, S. (2022). Leveraging IoT devices for improved healthcare accessibility in remote areas: an exploration of emerging trends. *Internet of Things and Edge Computing Journal*, *2*(1), 20-30

Ramjattan, N. A., Lala, V., Kousa, O., & Makaryus, A. N. (2017). Coronary CT angiography.

Silverglate, S. H. (2001). THE RESTATEMENT (THIRD) OF TORTS Products Liability—The Tension Between Product Design and... *Florida Bar Journal*, *75*(11).

Tobler, P., Cyriac, J., Kovacs, B. K., Hofmann, V., Sexauer, R., Paciolla, F., ... & Hirschmann, A. (2021). AI-based detection and classification of distal radius fractures using low-effort data labeling: evaluation of applicability and effect of training set size. *European Radiology*, *31*, 6816-6824.

Ueda, D., Kakinuma, T., Fujita, S., Kamagata, K., Fushimi, Y., Ito, R., ... & Naganawa, S. (2024). Fairness of artificial intelligence in healthcare: review and recommendations. *Japanese Journal of Radiology*, *42*(1), 3-15.

Verma, V., Chowdary, V., Gupta, M. K., & Mondal, A. K. (2018). IoT and robotics in healthcare. In *Medical big data and internet of medical things* (pp. 245-269). CRC Press.

Yang, H. Y., Hong, S. S., Yoon, J., Park, B., Yoon, Y., Choi, G. H., ... & Kim, S. H. (2024). Deep learning-based surgical phase recognition in laparoscopic cholecystectomy. *Annals of Hepato-Biliary-Pancreatic Surgery*, *28*(4), 466-473.

Zachariadis, C. B., & Leligou, H. C. (2024). Harnessing Artificial Intelligence for Automated Diagnosis. *Information*, *15*(6), 311.

Zaman, M. F., Elsayed, S. M., Ray, T., & Sarker, R. A. (2015). Evolutionary algorithms for dynamic economic dispatch problems. *IEEE Transactions on Power Systems*, *31*(2), 1486-1495.

Balancing Performance and Environmental Impact: Multi-Objective Lifecycle Optimization of Crumb Rubber-Modified Asphalt Binders

By Keshav Sekhri, Namitha Sudan, Neetu Gopakumar, and Krishna Prapoorna Biligiri

(Professor Department of Civil & Environmental Engineering, Indian Institute of Technology Tirupati, Andhra Pradesh)

Abstract

The major objective of this study was to optimize the asphalt binders modified with crumb rubber (CR) through a multi-objective lifecycle assessment approach by integrating performance enhancement with environmental sustainability through a lifecycle assessment (LCA) framework. Modified asphalt binders with varying CR (10, 15, and 20%) concentrations were prepared and subjected to standard physical tests, viscosity measurements, and rheological evaluation using a Dynamic Shear Rheometer. The LCA was conducted to quantify the environmental burdens across different modification levels, including energy use, emissions, and end-of-life impacts. Findings of the study showed that the CR modifier enhanced high-temperature performance and rutting resistance, i.e. CR improved elasticity and temperature susceptibility. Using a multi-objective decision-making framework, the optimal CR dosage was identified as 20%, representing the best compromise between performance enhancement and environmental benefit, with an 18.45% reduction in Global Warming Potential (GWP) compared to the virgin asphalt binder. The study highlighted the distinct role of CR in asphalt modification and supports their circular utilization through a comprehensive performance-environmental optimization approach, which not only extends pavement durability but also reduces carbon footprint and diverts substantial waste from landfills, offering a robust strategy for sustainable pavement engineering.

Keywords

Crumb Rubber, Asphalt-Rubber, Lifecycle Assessment, Multi-Objective Optimization, Sustainability.

Introduction

The global generation of end-of-life tires (ELT) has escalated in recent years due to rapid urbanization and economic growth. With billions of tires discarded annually, effective tire recycling has become a critical environmental as well as engineering priority. The improper ELT disposal techniques adopted mostly lead to land and water pollution, fire hazards, and long-term ecological degradation. To counter these threats, recycling and reuse of ELT has gained traction globally, with several valorization pathways being explored in civil infrastructure and transportation sectors. Through advanced processing technologies adopted at recycling plants, waste tires can be converted into reusable materials of varying sizes such as crumb rubber (CR), tire-derived aggregates (TDA), reclaimed rubber (RR), as well as micronized rubber powder (MRP), each offering diverse applications across geotechnical and transportation infrastructure realms.

Among these, the recyclability of CR is enhanced by its ease of large-scale production through mechanical shredding, rendering CR as a practical and sustainable material for extensive applications. CR has been successfully utilized in pavement systems through the development of asphalt-rubber (AR) technology since the early 1960s (McDonald 1981) through the wet process, where CR is pre-blended with the asphalt binder. Literature has established that CR-modified asphalt binders or AR binders contribute to long-term sustainability by reducing environmental impacts and diverting waste from landfills, which aligns strongly with the principles of the circular economy by converging both resource efficiency (Bartolozzi et al.; Bressi et al. 2021; Farina et al. 2017; Mohajerani et al. 2022; Sudan et al. 2025; Wang et al. 2018) as well as performance durability (Gawel et al. 2006; Harvey et al. 2001; Way et al. 2012) to support the development of resilient infrastructure.

Incorporating CR into asphalt binders has demonstrated significant improvements in performance, particularly in terms of thermal crack resistance, rutting resistance, and elasticity (Bressi et al. 2018; Gopakumar and Biligiri 2024a; b; Kaur and Biligiri 2025; Liao and Lo 2021; Nanjegowda and Biligiri 2020; Santagata et al. 2016; Venudharan et al. 2017). Although numerous studies have demonstrated the mechanical benefits of CR-modified asphalt binder, environmental impacts are often overlooked in traditional evaluations. Emerging studies have begun to address these concerns by adopting lifecycle assessment (LCA) frameworks to evaluate the environmental burdens of recycled rubber products. For instance, a recent study evaluated the potential environmental impacts of producing CR, MRP, and RR under scenarios involving a transition to renewable resources in the energy sector (Sudan et al. 2025). The LCA findings indicated that CR consistently exhibited the lowest global warming potential (GWP), regardless of the ELT source or electricity consumption. However, integrating performance parameters and environmental indicators are crucial to further understand the overall sustainability benefits of the recycled materials utilization in constructing high-performing resilient infrastructure.

In this direction, a recent study proposed an integrated framework that combined Balanced Mix Design (BMD) and LCA of conventional asphalt mixtures, which demonstrated improved rutting and cracking resistance, while significantly reduced GWP, respiratory effect, ecotoxicity, and life-cycle costs (Wang et al. 2025). Another study developed a lifecycle-based optimization framework for Warm Mix Asphalt (WMA) incorporating recycled materials such as waste tire rubber and quantified the environmental impacts during production and also identified sustainable material blends using a multi-objective approach (Polo-Mendoza et al. 2023). In the context of CR-modified asphalt binders or AR, there is a critical need to optimize asphalt binder formulations not just for mechanical performance, but also for environmental sustainability. Given that CR-modified asphalt binders vary in both their rheological behavior and lifecycle impacts depending on dosage levels, a data-driven multi-objective optimization framework is essential, which would enable stakeholders to adopt asphalt binder formulations that offer the best material efficiency. Further, the literature on the LCA of virgin asphalt production from the refinery without any additives and its comparison with the asphalt binder modified with the

ground tire rubber, it was observed that crude oil extraction was the primary driver of impacts ("Life Cycle Assessment of Asphalt Binder - Asphalt Institute" 2021).

Therefore, the major objective of this study was to assess the correlations of environmental impact indicators along with fundamental consistency and rheological performance of AR binders through multi-objective optimization of CR dosage by adopting lifecycle assessment approach. The scope of the study (Figure 1) included: (i) preparation of viscosity graded VG-30 asphalt binder with three CR dosages: 10, 15, and 20% by weight of asphalt binder, (ii) fundamental consistency tests to assess performance of virgin and CR-modified asphalt binder blends, (iii) lifecycle inventory (LCI) analysis of raw material production and CR recycling processes, (iv) LCA of virgin and CR-modified asphalt binder blends, (v) correlation analysis of CR content and corresponding environmental and performance parameters, and (vi) multi-objective optimization of optimal CR dosage, balancing performance enhancement with environmental sustainability. The current work builds upon and extends the methodology developed in our previous study Sudan et al. (2025). It is envisioned that the findings of this study will support the strategic reuse of ELT in high-performance, low-carbon pavement systems, advancing long-term goals of resource circularity and sustainable infrastructure.

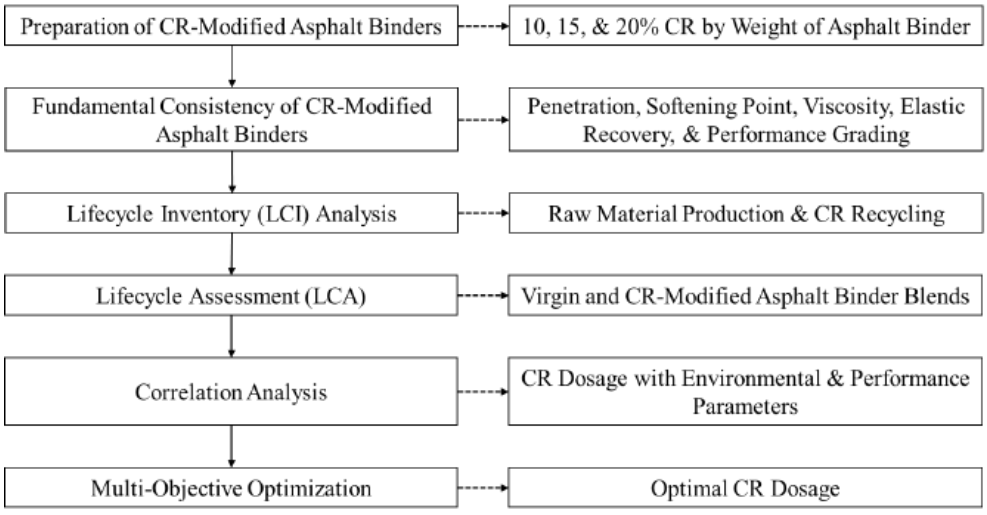


Figure 1: Research Outline

1.Methodology

This study aims to evaluate the environmental performance of CR-modified asphalt binder using an LCA-based multi-objective optimization framework. The methodology adopted in this study comprised three integrated stages: laboratory experimentation, LCA, and multi-objective optimization. In the first stage, modified asphalt binders were prepared by incorporating varying proportions of CR. Standard laboratory tests were conducted to evaluate their fundamental consistency and rheological properties such as Penetration, softening point, ductility, –viscosity, and performance grading to assess the functional suitability of the modified

asphalt binders for paving applications. In the second stage, a cradle-to-gate LCA was performed to quantify the environmental impacts associated with each asphalt binder formulation and correlation analysis was carried out to understand the trade-offs between asphalt binder performance parameters and environmental impact indicators.

1.1 Raw Materials and Asphalt Binder Tests

The AR blends were blended through the wet process, a method commonly known to be effective in adding CR to asphalt binder matrices, which allows for maximum dispersion and interaction between the base asphalt binder and the rubber modifier, thus improving the elasticity and thermal stability of the resulting blend. Two reference base binders, AC10 and AC30, were employed in AR blend formulations to evaluate their relative performance when compared with CR modification. The virgin asphalt binder was first heated in a stainless-steel tank to a temperature of around 160 °C to obtain pourable consistency. After the asphalt binder reached the optimum fluidity, pre-weighed CR particles (CR30 grade, with maximum particle size 600 μm) were slowly introduced into the hot binder. The blending was accomplished under conditions of high shear mixing for uniformity and complete interaction. The shear rate was held constant at 1000 rpm, and the temperature was increased to 180 °C. The high-temperature mixing was continued for 60 min to swell the CR and develop a uniform, gel-like modified binder.

Three different AR binder mixtures were prepared for each of the two base binders, AC10 and AC30, corresponding to dosages of CR at 10, 15, and 20% by weight of binder. These samples had standardized notations to provide consistency in the study. The asphalt binder blends that utilized the AC10 base binder were designated as 10CR10, 10CR15, and 10CR20, indicating AC10 binder blended with 10, 15, and 20% CR, respectively by weight of the asphalt binder. Similarly, blends with the base binder AC30 were designated as 30CR10, 30CR15, and 30CR20, with CR dosages of 10, 15, and 20% by weight of the asphalt binder. Figure 2 shows the fundamental asphalt binder tests such as elastic recovery, Penetration test (ASTM D5 1997), ductility (ASTM D113 1999), and softening point (ASTM D36 2014) evaluations.

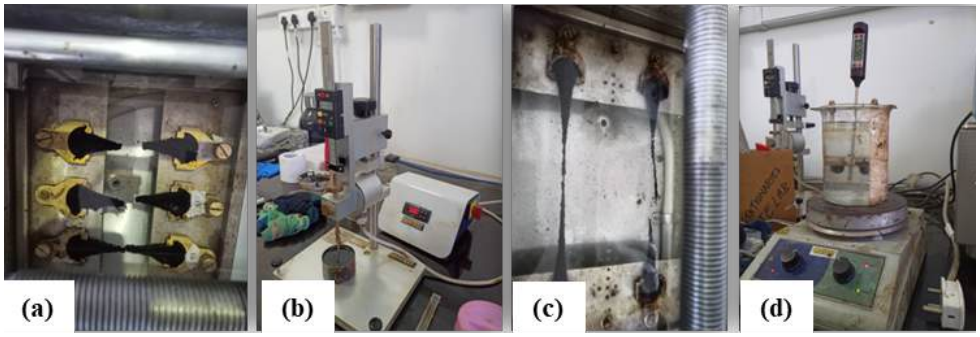


Figure 2: Fundamental Asphalt Binder Tests: (a) Elastic Recovery, (b) Penetration Test (c) Ductility Test, and (d) Softening Point Test

1. 2. Environmental Impact Assessment

1.2.1 Goal & Scope

In India, about 0.27 million ELTs/day are discarded within the country while 3 million ELTs/day are imported to generate recycled rubber products that are used as secondary raw materials due to high demand of recycled goods such as CR, promoting resourceful consumption in various industries including pavement engineering (Gaidhane et al. 2022).

The incorporation of CR, derived from ELTs, presents a promising and sustainable solution due to its low environmental footprint and potential for beneficial reuse in asphalt binder modification and AR pavement applications. The results from the previous study on the environmental impact of ELT recycling process indicated that CR had the lowest GWP, irrespective of ELT source and electricity scenarios, while the production of other similar products such as micronized rubber powder and reclaimed rubber would lead to higher environmental impact (Sudan et al. 2025).

Thus, the goal of this study was to quantify the environmental impacts associated with the production of CR-modified asphalt binder at the manufacturing plant level, using a cradle-to-gate lifecycle assessment approach. The analysis aims to support informed decision-making regarding material selection and process optimization in pavement construction, with a focus on enhancing sustainability outcomes. The recycling industry is heavily dependent on the electricity sources, where the country is planning to shift 50% cumulative electric power installed capacity from non-fossil fuel energy sources to renewable resources by 2030. Background data and inventory information for CR manufacturing in recycling plants was retained from the prior analysis, ensuring consistency and comparability across the two studies. The study used a declared unit of 1 metric ton production of asphalt binder blended with CR and virgin asphalt binder. The scope of the study encompassed a cradle-to-gate system boundary, which included raw material extraction, CR manufacturing, and CR-modified asphalt binder production.

1.2.2 Lifecycle Inventory Analysis

Lifecycle inventory (LCI) analysis is the second phase of LCA involving the compilation and quantification of inputs and outputs of the product systems. In this study, the product system was modeled using OpenLCA© GreenDelta software with Ecoinvent™ v3.10 as LCI database (“ecoinvent database” 2024). The primary data of ELT recycling and blending plants was specific to the plant operation processes and over which the practitioner had direct control. Further, this study used characterization factors of ReCiPe 2016 lifecycle impact assessment (LCIA) method at midpoint level, which strongly correlated with the environmental flows (Huijbregts et al. 2017). Table 1 represents the energy consumption of a blending plant. The background electricity data for both plants in this study was adopted from Scenario S6 reported by (Sudan et al. 2025), which reflected a balanced energy mix wherein 50% of the conventional sources were replaced by renewable energy sources.

Table 1: Energy Consumption of Blending Plant

Equipment in Blending Plant	Power Consumption	Unit
High Shear Colloidal Mixing Unit	30	kW
Asphalt Pump	9	kW
Unit Convertor	4	kW
Control Unit	2	kW
Total Power Consumption	45	kW

1.3. Environmental Optimization

In this study, a reinforcement learning (RL) based decision-making framework using Q-learning was used as a multi-objective optimization framework to select the most sustainable and effective binder blend. Environmental indicators are categorized under human health, ecosystem quality, and resource depletion, while performances including softening point, upper PG (unaged), and Upper PG (RTFO-aged) were also integrated. Figure 3 illustrates the Q-learning based decision framework developed for the identification of optimal CR-modified asphalt binder blends.

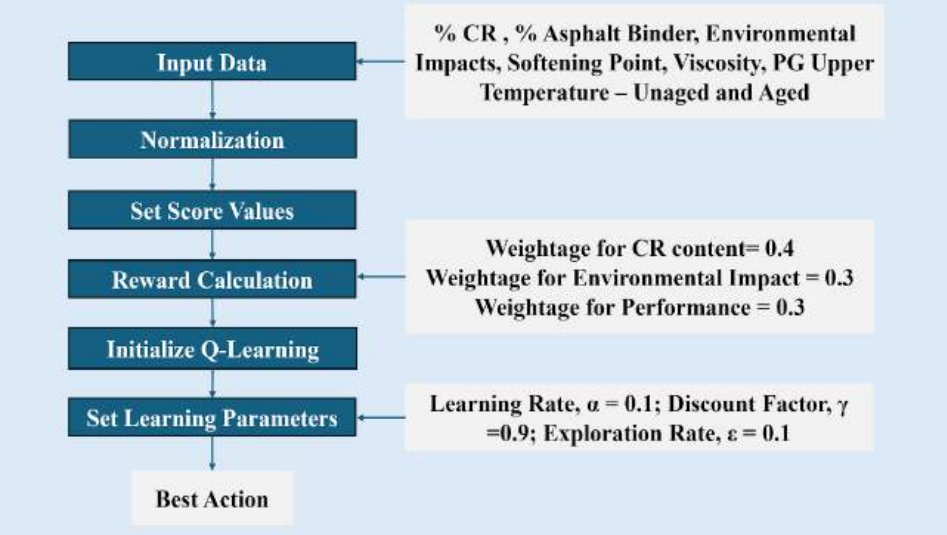


Figure 3: Q-learning based decision framework

1.3.1 Data Preparation and Indicator Classification

The dataset used in this study comprises eight binder blends with varying CR contents (10, 15, and 20%) and two binder grades (AC10 and AC30). The environmental indicators were classified into three categories: (i) human health impacts - including GWP, Human Toxicity Potential (cancerous and non-cancerous), Ionizing Radiation Potential (IRP), Ozone Depletion Potential (ODP), Particulate Matter Formation Potential (PMFP), Photochemical Oxidant Formation Potential affecting humans (HOFP), and Water Consumption Potential (WCP, which partially affects human health); (ii) ecosystem impacts - including GWP, Terrestrial Acidification Potential (TAP), Freshwater and Marine Eutrophication Potentials (FEP, MEP), Ecotoxicity

Potentials (FETP, METP, TETP), Agricultural Land Occupation Potential (LOP), Photochemical Oxidant Formation Potential affecting ecosystems (EOFP), and WCP (also partially affecting ecosystems); and (iii) resource depletion impacts - represented by Fossil Fuel Potential (FFP) and Surplus Ore Potential (SOP). All input parameters were normalized using min-max scaling to eliminate dimensional bias and ensure comparability among indicators. The normalization facilitated a balanced reward computation, preventing any single indicator from disproportionately influencing the learning process due to scale differences. An RL approach based on the Q-learning algorithm was employed to identify the mix that offers the optimal trade-off across all considered objectives. The environment was modeled with each mix configuration representing both a state and a potential action.

1.3.2 Reward Function Formulation

A composite reward function was developed to understand the trade-offs between environmental sustainability, resource efficiency, and asphalt binder performance. The reward function was structured as a weighted sum of normalized variables. Specifically, a weight of 0.4 was assigned to CR content to encourage its maximization, while 0.3 was assigned to GWP due to its combined influence on human health and ecosystem categories. The remaining environmental categories - human health, ecosystem, and resource depletion, were each assigned a weight of 0.1, with a target to minimize. Additionally, a total weight of 0.3 was allocated to performance parameters, distributed equally among the three indicators, all of which were intended to be maximized to facilitate the selection of asphalt binders exhibiting superior high-temperature performance.

1.3.3 Q-Learning Implementation for Optimization

A reinforcement learning approach based on the Q-learning algorithm was employed to identify the blend that offers the optimal trade-off across all considered objectives. Q-learning is a model-free reinforcement learning technique, which iteratively learns optimal actions that maximize a predefined reward function. In this study, each asphalt binder blend was treated as both a state and an action within the environment. The Q-table, initialized as a zero matrix, was updated over 1000 iterations using standard Q-learning update rules with a learning rate (α) of 0.1, discount factor (γ) of 0.9, and an exploration rate (ε) of 0.1. The Q-learning algorithm updates the value of each state–action pair using Equation 1. The action with the highest Q-value at the conclusion of training was selected as the optimal blend of the asphalt binder and CR.

$$Q(S, A) = Q(S, A) + \alpha * [\text{reward} + \gamma * \max(Q(S', \text{all actions})) - Q(S, A)] \quad (1)$$

where Q(S,A) represents the estimated value of taking action A in state S, α = is the learning rate, γ is the discount factor, S' is the next state.

2.Results

2.1 Laboratory Tests

The properties of asphalt binders modified with varying percentages of CR demonstrated notable improvements in high-temperature performance, accompanied by reductions in

Penetration and ductility as shown in the Table 2. A progressive decrease in penetration values was observed with increasing CR content, indicating enhanced binder stiffness. Specifically, the penetration of the AC10-based binders decreased from 128.8 (unmodified) to 41.8 at 20% CR, while for AC30-based binders, the corresponding reduction was from 67.36 to 21.91. Further, it was noted that the ductility also declined with higher CR content, reflecting increased elasticity and reduced flow characteristics. In contrast, the softening point exhibited a consistent increase with CR modification, indicating improved thermal resistance. The softening point for AC10-based binders increased from 38.83 °C (unmodified) to 55.15 °C at 20% CR, and from 46.8°C to 65.65°C for AC30-based binders. The enhancement in softening point is indicative of better high-temperature stability. Correspondingly, the performance grading (PG) values also improved. The upper PG temperature for unaged binders rose from 64 to 82 °C, and the RTFO-aged binders exhibited an increase from 64 °C/70 °C (for unmodified AC10/AC30) to 88 °C at 20% CR. The viscosity results clearly demonstrated that the addition of CR significantly increased the viscosity of asphalt binders across all temperatures tested (120, 135, and 150 °C). Compared to the unmodified binders, CR-modified blends - particularly with 20% CR exhibited a substantial increase in viscosity, indicating enhanced stiffness. Figure 4 displays the ASTM A_i-VTS_i curve illustrating the relationship between a binder viscosity and temperature, which indicated that the inclusion of CR substantially improved the high-temperature rheological behavior of asphalt binders, rendering them more suitable for resisting rutting in hot climate conditions.

Table 2: Fundamental asphalt binder test results and PG grades

Mix	Penetration (0.1 mm)	Ductility (cm)	Softening Point (°C)	Upper PG (unaged) (°C)	Upper PG (RTFO aged) (°C)
AC10	128.8	100	38.83	64	64
10AC10	114.8	13.05	46.75	76	76
15AC10	51.5	14.05	48.1	82	88
20AC10	41.8	8.45	55.15	82	88
AC30	67.36	10	46.8	70	70
10AC30	22.58	14.6	51.45	76	82
15AC30	27.08	12.15	56	82	88
20AC30	21.91	9.95	65.65	82	88

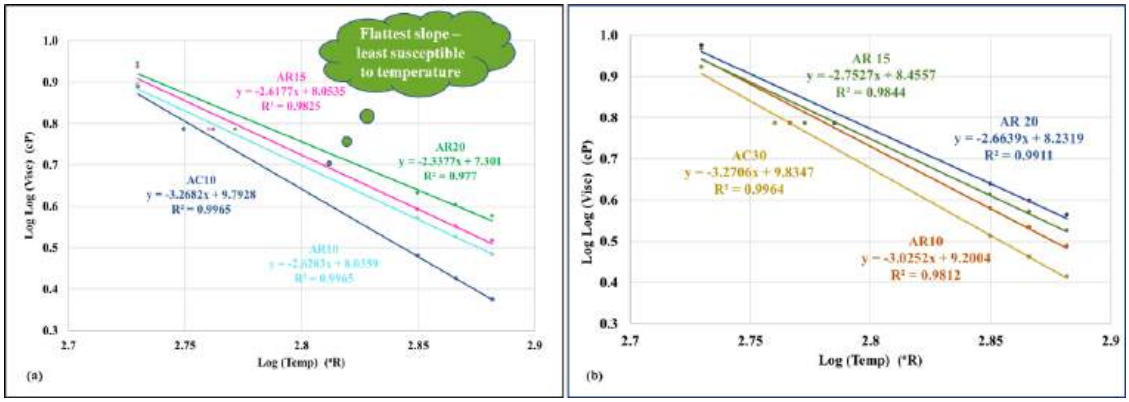


Figure 4: ASTM A_i-VTS_i curves for (a) AC10 and CR-modified AC10 binders, and (b) AC30 and CR-modified AC30 binder

2.2 Environmental Impact Assessment

Table 3 shows the environmental impact assessment, which revealed that the CR-modified asphalt binder blends consistently exhibited lower values across all impact categories when compared to virgin asphalt binder. The virgin binder exhibited the highest values across most environmental indicators, including a GWP of 889.44 kg CO_2-Eq, Freshwater Ecotoxicity Potential (FETP) of 10.45 kg 1,4-DCB eq, and Fossil Fuel Potential (FFP) of 1133.47 MJ, establishing it as the least sustainable option among those evaluated. As the CR content increased from 0 to 20%, a significant reduction in GWP was observed, with the lowest value of 724.82 kg CO_2-Eq. recorded for the 80:20 (asphalt binder: CR) blend, indicating a 18.45% reduction in GWP. The highest percentage reduction was observed in EOFP, which decreased by approximately 19.11%, followed by PMFP with 18.93%, and HOFP at 18.89%. Additionally, notable reductions included HTPc at 18.72%, FFP at 19.48%, and TETP at 19.57%. Conversely, the lowest percentage reduction was observed in ODP, which marginally decreased by only 5.26%. The limited improvement was likely because such global-scale indicators were primarily influenced by upstream industrial processes and emissions, which are less affected by the substitution of crumb rubber for virgin asphalt binder. Additionally, indicators such as FEP, associated with water - related impacts tend to exhibit minimal changes, as they are governed more by overall production processes than by the specific binder materials used.

Table 3: LCA results of asphalt binder blends

Impact Indicator	Unit	10 % CR	15 % CR	20 % CR	Virgin Binder
TAP	kg SO_2-Eq	2.12	2.01	1.91	2.34
GWP	kg CO_2-Eq	807	766	725	889
FETP	kg 1,4-DCB-Eq	9.62	9.20	8.79	10.4
METP	kg 1,4-DCB-Eq	20.31	19.35	18.39	22.2
TETP	kg 1,4-DCB-Eq	6914	6540	6165	7663
FFP	kg oil-Eq	1022	967	912	1133
FEP	kg P-Eq	0.056	0.055	0.054	0.057
MEP	kg N-Eq	0.023	0.022	0.021	0.025
HTPc	kg 1,4-DCB-Eq	103.95	98.59	93.23	114.68
HTPnc	kg 1,4-DCB-Eq	206.52	199.97	193.43	219.62
IRP	kBq Co-60-Eq	6.95	6.86	6.78	7.12
LOP	m^2*a crop-Eq	6.78	6.46	6.14	7.42
SOP	kg Cu-Eq	10.49	9.92	9.35	11.62
ODP	kg CFC-11-Eq	0.00019	0.00018	0.00017	0.00021
PMFP	kg PM2.5-Eq	0.79	0.75	0.72	0.86
HOFP	kg NO_x-Eq	2.20	2.08	1.97	2.43
EOFP	kg NO_x-Eq	2.84	2.69	2.54	3.14
WCP	m^3	1.21	1.17	1.13	1.29

Lowest Impact		Highest Impact	

2.3 Correlation Analysis

The correlation analysis, as illustrated in Figure 5, reveals distinct relationships between binder composition, performance, and GWP. A strong inverse correlation (r = - 1.00) between CR content and GWP indicates that increasing CR significantly reduces emissions, while a perfect positive correlation with binder content (r = 1.00) confirms higher emissions with more virgin binder. Furthermore, penetration, reflecting binder softness, showed a moderate to strong positive correlation with GWP (r = 0.6627), suggesting that softer binders may be less sustainable. CR content exhibited strong positive correlations with high-temperature performance indicators, including viscosity at 135°C (r = 0.8014), softening point (r = 0.815), and upper PG grades for both unaged (r = 0.9428) and RTFO-aged binders (r = 0.9429). The penetration showed strong negative correlations with softening point and viscosity (r = -0.8102 and -0.6528, respectively), indicating that as binders become stiffer, they become less penetrable. The strong interrelations among softening point, viscosity, and PG grades ranging from 0.7096 to 0.983 affirm that combined significance in evaluating binder performance is important.

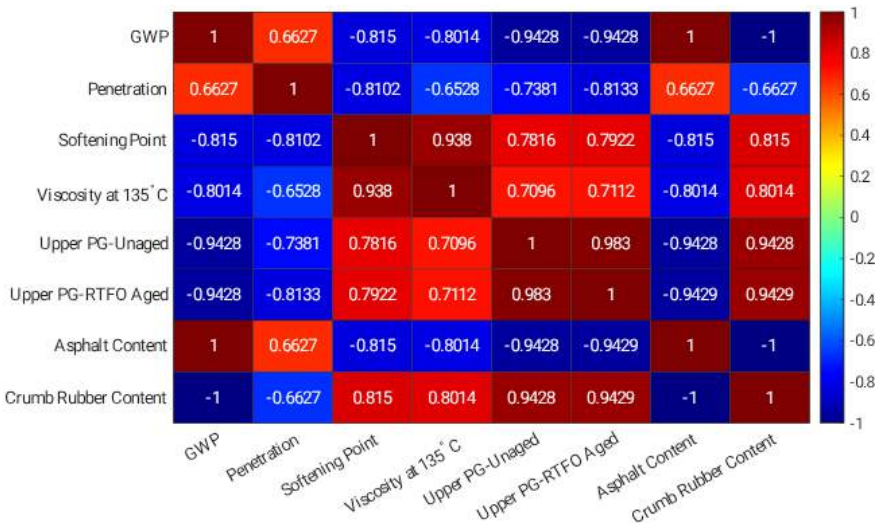


Figure 5: Correlation matrix between rheological parameters

To evaluate the trade-offs and potential burden shifting among environmental impact categories, a correlation heatmap was developed as shown in Figure 6. The results demonstrated strong positive correlations ($r \approx 1$) among most indicators, including GWP, FETP, METP, and TETP, suggesting that improvements in one category are likely to yield concurrent benefits in others. CR content exhibited a strong inverse correlation ($r = -1$) with all major impact categories, indicating its effectiveness in mitigating overall environmental burdens. In contrast, the asphalt binder content showed perfect positive correlations with all indicators, reaffirming its substantial contribution to environmental impacts. Notably, FEP presented slightly weaker correlations with the other categories, implying the possibility of burden shifting which may be attributed to its distinct sensitivity to electricity consumption during asphalt binder production.

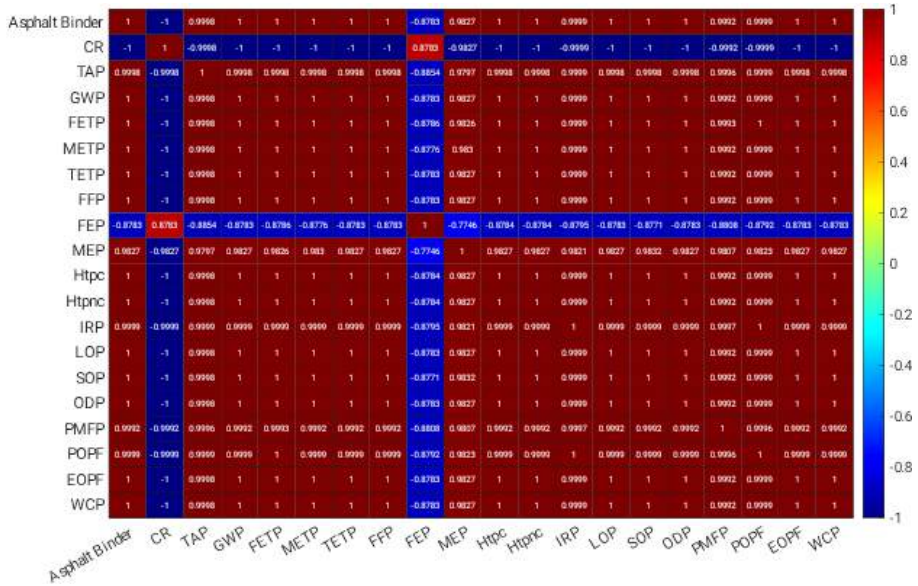


Figure 6: Heatmap showing correlations between binder properties and environmental impacts, indicating trade-offs and potential burden shift

2.4 Optimization

The optimized asphalt binder blend incorporating 20% CR with AC10 binder demonstrated a favorable balance between mechanical performance and environmental impact. The binder exhibited a softening point of 55.15 °C, with upper performance grades of 82 °C (unaged) and 88 °C (RTFO-aged), indicating improved resistance to rutting at higher temperatures, with least environmental impact. Table 4 shows the result of a developed multi-objective optimization framework.

Table 4: Results of Q-Learning Optimization

Mix ID	20AC10
Binder Type	AC10
Softening Point (°C)	55.15
Upper PG (unaged) (°C)	82
Upper PG (RTFO aged) (°C)	88
GWP	724.82
HTPc	93.23
HTPnc	193.43
IRP	6.78
ODP	0.00017
PMFP	0.72
WCP	1.13
TAP	1.91
FEP	0.057
MEP	0.021
FETP	8.79
METP	18.39
TETP	6165.4
LOP	6.14
EOFP	2.54
WCP	1.13
FFP	912.14
SOP	9.35

3.Conclusion and Recommendation

The major objective of this study was to assess the correlations of environmental impact indicators along with fundamental consistency and rheological performance of AR binders through multi-objective optimization of CR dosage by adopting lifecycle assessment approach. The study developed a novel decision-making framework integrating LCA and Q-learning based multi-objective optimization to identify optimal CR-modified asphalt binder composition based on environmental and performance indicators. The results demonstrated that increasing CR content consistently reduced GWP and most environmental impact categories, with the highest percentage reduction observed in EOFP (-19.11%) and the lowest in ODP (-5.26%). It has to be noted that incorporation of 20% CR in the asphalt binder has reduced the GWP by 18.45%.

Further, the CR content also showed a strong positive correlation with high-temperature properties such as softening point and viscosity, highlighting its effectiveness in enhancing binder performance. However, LCA results revealed that indicators such as FEP showed weaker correlation, reinforcing the influence of background processes such as electricity mix on these categories.

Additionally, the Q-learning framework used in this study enabled single-solution selection rather than traditional pareto optimization, ensuring convergence toward a blend that best satisfies both performance and sustainability objectives. The integration of reinforcement learning with LCA enables a dynamic exploration of trade-offs among conflicting objectives, offering a robust approach to materials selection in sustainable pavement engineering. The proposed methodology not only builds upon but also extends the framework established in prior work by introducing a data-driven optimization approach that supports environmentally responsible and performance-optimized decision-making in asphalt binder selection.

Overall, this study highlights the benefit of CR-modified asphalt binder in reducing environmental burdens and enhancing binder performance. Future work may focus on extending the scope to full-scale AR pavement systems providing more comprehensive insights into performance and environmental sustainability. Additionally, mix design optimization for AR pavements should consider a broader range of rheological and mechanistic properties beyond which are explored in this study.

Acknowledgement

The authors would like to thank the Tinna Rubber and Infrastructure Limited, India for providing the inventory data. The authors are also thankful to Hindustan Colas India and Tinna Rubber and Infrastructure Limited, India for providing the raw materials.

Works Cited

ASTM D5. 1997. “ASTM International. (1997). ASTM D5–97: Standard test method for penetration of bituminous materials. West Conshohocken, PA: ASTM International.”

ASTM D36. 2014. “ASTM International. (2014). Standard test method for softening point of bitumen (ring-and-ball apparatus) (ASTM D36/D36M − 14). West Conshohocken, PA: ASTM International. https://doi.org/10.1520/D0036_D0036M-14.”

ASTM D113. 1999. “ASTM International. (1999). Standard test method for ductility of bituminous materials (ASTM D113 – 99). West Conshohocken, PA: ASTM International.”

Bartolozzi, I., I. Antunes, and F. Rizzi. n.d. “The environmental impact assessment of Asphalt Rubber: Life Cycle Assessment.”

Bressi, S., N. Colinas-Armijo, and G. Di Mino. 2018. “Analytical approach for the mix design optimisation of bituminous mixtures with crumb rubber.” Mater Struct, 51 (1). Springer Science and Business Media LLC. https://doi.org/10.1617/s11527-018-1152-9.

Bressi, S., J. Santos, M. Orešković, and M. Losa. 2021. “A comparative environmental impact analysis of asphalt mixtures containing crumb rubber and reclaimed asphalt pavement using life cycle assessment.” International Journal of Pavement Engineering, 22 (4): 524–538. Informa UK Limited. https://doi.org/10.1080/10298436.2019.1623404.

“Ecoinvent database.” 2024. ecoinvent. Accessed August 1, 2024. https://ecoquery.ecoinvent.org/3.10/cutoff/search.

Farina, A., M. C. Zanetti, E. Santagata, and G. A. Blengini. 2017. “Life cycle assessment applied to bituminous mixtures containing recycled materials: Crumb rubber and reclaimed asphalt pavement.” Resources, Conservation and Recycling, 117: 204–212. Elsevier BV. https://doi.org/10.1016/j.resconrec.2016.10.015.

Gaidhane, J., I. Ullah, and A. Khalatkar. 2022. “Tyre remanufacturing: A brief review.” Materials Today: Proceedings, International Conference on Latest Developments in Materials & Manufacturing, 60: 2257–2261. https://doi.org/10.1016/j.matpr.2022.04.142.

Gawel, I., R. Stepkowski, and F. Czechowski. 2006. “Molecular Interactions between Rubber and Asphalt.” Ind. Eng. Chem. Res., 45 (9): 3044–3049. American Chemical Society (ACS). https://doi.org/10.1021/ie050905r.

Gopakumar, N., and K. P. Biligiri. 2024a. “Morphological and Rheological Assessment of Waste Plastic-Modified Asphalt-Rubber Binder.” 405–415. Springer.

Gopakumar, N., and K. P. Biligiri. 2024b. “Can rubberized-asphalt comprising tire-derived aggregates produced through dry process be an alternative subsurface road material?” Construction and Building Materials, 439: 137270. Elsevier BV. https://doi.org/10.1016/j.conbuildmat.2024.137270.

Harvey, J., M. Bejarano, and L. Popescu. 2001. “Accelerated Pavement Testing of Rutting and Cracking Performance of Asphalt-Rubber and Conventional Asphalt Concrete Overlay Strategies.” Road Materials and Pavement Design, 2 (3): 229–262. Informa UK Limited. https://doi.org/10.1080/14680629.2001.9689902.

Huijbregts, M. A. J., Z. J. N. Steinmann, P. M. F. Elshout, G. Stam, F. Verones, M. Vieira, M. Zijp, A. Hollander, and R. van Zelm. 2017. “ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level.” Int J Life Cycle Assess, 22 (2): 138–147. https://doi.org/10.1007/s11367-016-1246-y.

Kaur, H., and K. P. Biligiri. 2025. “Advanced thermochemical characterization and environmental assessment of modified asphalt-rubber incorporated with industrial wastes: A novel strategy to promote circular economy in sustainable road construction.” Resources, Conservation and Recycling, 218: 108258. Elsevier BV. https://doi.org/10.1016/j.resconrec.2025.108258.

Liao, M.-C., and T.-J. Lo. 2021. “Material Characterization and Balanced Design of Asphalt–Rubber Binders.” J. Mater. Civ. Eng., 33 (1). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/(asce)mt.1943-5533.0003504.

“Life Cycle Assessment of Asphalt Binder - Asphalt Institute.” 2021. Accessed July 24, 2025. https://www.asphaltinstitute.org/engineering/sustainability/life-cycle-assessment-of-asphalt-binder/.

McDonald, C. 1981. “Recollections of early asphalt-rubber history.” 27–29. US Department of Transportation Washington, DC, USA.

Mohajerani, A., H. Kurmus, D. Conti, L. Cash, A. Semcesen, M. Abdurahman, and M. T. Rahman. 2022. “Environmental impacts and leachate analysis of waste rubber incorporated in construction and road materials: A review.” Science of The Total Environment, 835: 155269. Elsevier BV. https://doi.org/10.1016/j.scitotenv.2022.155269.

Nanjegowda, V. H., and K. P. Biligiri. 2020. “Recyclability of rubber in asphalt roadway systems: A review of applied research and advancement in technology.” Resources, Conservation and Recycling, 155: 104655. Elsevier BV. https://doi.org/10.1016/j.resconrec.2019.104655.

Polo-Mendoza, R., G. Martinez-Arguelles, and R. Peñabaena-Niebles. 2023. “A multi-objective optimization based on genetic algorithms for the sustainable design of Warm Mix Asphalt (WMA).” International Journal of Pavement Engineering, 24 (2). Informa UK Limited. https://doi.org/10.1080/10298436.2022.2074417.

Santagata, E., M. Lanotte, O. Baglieri, D. Dalmazzo, and M. C. Zanetti. 2016. “Analysis of bitumen–crumb rubber affinity for the formulation of rubberized dry mixtures.” Mater Struct, 49 (5): 1947–1954. Springer Science and Business Media LLC. https://doi.org/10.1617/s11527-015-0625-3.

Sudan, N., A. Singh, C. G. Bhat, and K. P. Biligiri. 2025. “Lifecycle assessment of end-of-life tire recycling through pathways of transitioning the recycling industry to renewable energy sources.” Resources, Conservation and Recycling, 216: 108169. Elsevier BV. https://doi.org/10.1016/j.resconrec.2025.108169.

Venudharan, V., K. P. Biligiri, J. B. Sousa, and G. B. Way. 2017. “Asphalt-rubber gap-graded mixture design practices: a state-of-the-art research review and future perspective.” Road

Materials and Pavement Design, 18 (3): 730–752. Informa UK Limited. https://doi.org/10.1080/14680629.2016.1182060.

Wang, J., R. Zhang, H. Zhou, W. Huang, D. Feng, and X. Li. 2025. "Optimization of asphalt mix design considering mixture performance, environmental impact, and life cycle cost." Journal of Cleaner Production, 512: 145618. https://doi.org/10.1016/j.jclepro.2025.145618.

Wang, T., F. Xiao, X. Zhu, B. Huang, J. Wang, and S. Amirkhanian. 2018. "Energy consumption and environmental impact of rubberized asphalt pavement." Journal of Cleaner Production, 180: 139–158. Elsevier BV. https://doi.org/10.1016/j.jclepro.2018.01.086.

Way, G. B., K. Kaloush, and K. P. Biligiri. 2012. "Asphalt-rubber standard practice guide." Tempe: Rubber Pavements Association.

Most Likely Resting Place of MH370 By Vihaan Manik



MH 370 was scheduled to depart Kuala Lumpur International Airport (WMKK) in Malaysia on the morning of the 8th of March, 2014. In command was Captain Zaharie Ahmad Shah, a highly experienced pilot logging more than 8,000 flight hours on the Boeing 777, the variant of aircraft being used for this flight. Alongside him, as First Officer, was Flight Cadet Fariq Abdul Ahmed. He was new to this variant of aircraft, and was put alongside Captain Zaharie to make his switchover from the Airbus A330 to the Boeing 777. The specific aircraft used for this approximately 5h 30m flight from Kuala Lumpur (WMKK) to Beijing Capital (ZBAA) in China was a nearly 16-year-old Boeing 777-2H6ER, registered in Malaysia as 9M-MRO (Niner Mike - Mike Romeo Oscar). At the time of the flight, the aircraft was reported to be in perfect working order, with the only maintenance being done pre-flight was topping up the flight crew's oxygen cylinders, which is routine maintenance.

The flight crew decided to fill the plane up with 49,100 kilograms of fuel for the journey, which would be enough for approximately 7h 35m of flying time, but was necessary to overfuel in case of emergency. Precisely 42 minutes before the plane departed Kuala Lumpur, it logged on to the Inmarsat Satellite Communication Network. Onboard were 227 passengers and 12 crew, from a combined total of 14 nations. 2 of the passengers onboard were later found to be using fake passports, and actually having Iranian citizenship, but were ruled out of any wrongdoings since they were most probably seeking refugee status in China. Out of the 239 souls on board, 153 of them were Chinese, while 50 were Malaysian. This was due to the flight being between these two nations, and had nothing abnormal in it.

At precisely 00:27 MST (UTC + 8), MH 370 was provided clearance from WMKK ATC for push and start. Approximately 13 minutes after that, ATC provided clearance for take-off from Runway 32R, and MH 370 was seen for the last time one minute later. 53 seconds after take-off, MH 370 is provided instructions from the ATC to climb to FL180 (18,000 ft) and

proceed towards waypoint IGARI. 3 minutes after this, the aircraft is transferred from Kuala Lumpur Airport ATC to Kuala Lumpur Radar ATC, since it has now left the vicinity of the airport. This is followed by clearance to climb to FL250, and then, a few minutes later, to FL350, which would be its cruising altitude for the flight. At 01:01 local time, around 11 minutes after the last call, the Captain informs the ATC that the plane is now cruising at FL350. While this is not a required call, it is still used by Captains often to inform the ATC about their relative altitude. At 01:07 local, the final data transmission using ACARS (Aircraft Communication Addressing and Responding System) is made. A few seconds after that, Captain Zaharie informs the ATC again about their altitude, FL350.

Now, a pilot normally only has two reasons to reiterate what they have said to the ATC. The first, the much simpler and common one, is when they simply forgot if they have made the call and want to reconfirm. But this is not a possibility in this scenario, since the call Captain Zaharie repeated was not a necessary one to make. The second possibility, the complex one, is when a pilot has been away from that radio frequency for a certain amount of time, most probably on another radio frequency, and wants to check with the ATC if there were any calls they missed. This is the concerning part. As the plane was on another radio frequency for the whole of 7 minutes, there is a very high possibility that it was communicating to another organisation, likely a terrorist organisation for this time period. Another thing worth pointing out is that when Captain Zaharie stated the repeated call, he sounded a lot more stressed, as if he was working on something else while speaking to ATC.

At 01:19 local, the plane is handed over from Kuala Lumpur Radar Control to Ho Chi Minh Radar Control, and this is where things start getting heated. As the aircraft passes waypoint IGARI, the Captain makes one final call to ATC- “Good night. Malaysian three seven zero.” At this point, the plane was passing from the Malaysian fir to the Vietnamese fir, and was out of ATC contact for a few minutes. This is when the pilots shut off autopilot and the transponder, taking a sharp turn heading South-South-West. Nearby, a Thai military defence zone was monitoring the aircraft, but thought nothing of the turn, thinking it was a regular diversion of a Malaysian aircraft back to Malaysia. This is where it begins to seem like this was all planned. The turn was done at an angle higher than 40°, which is not possible with autopilot on, and was done so that the Thai military airspace can be avoided.

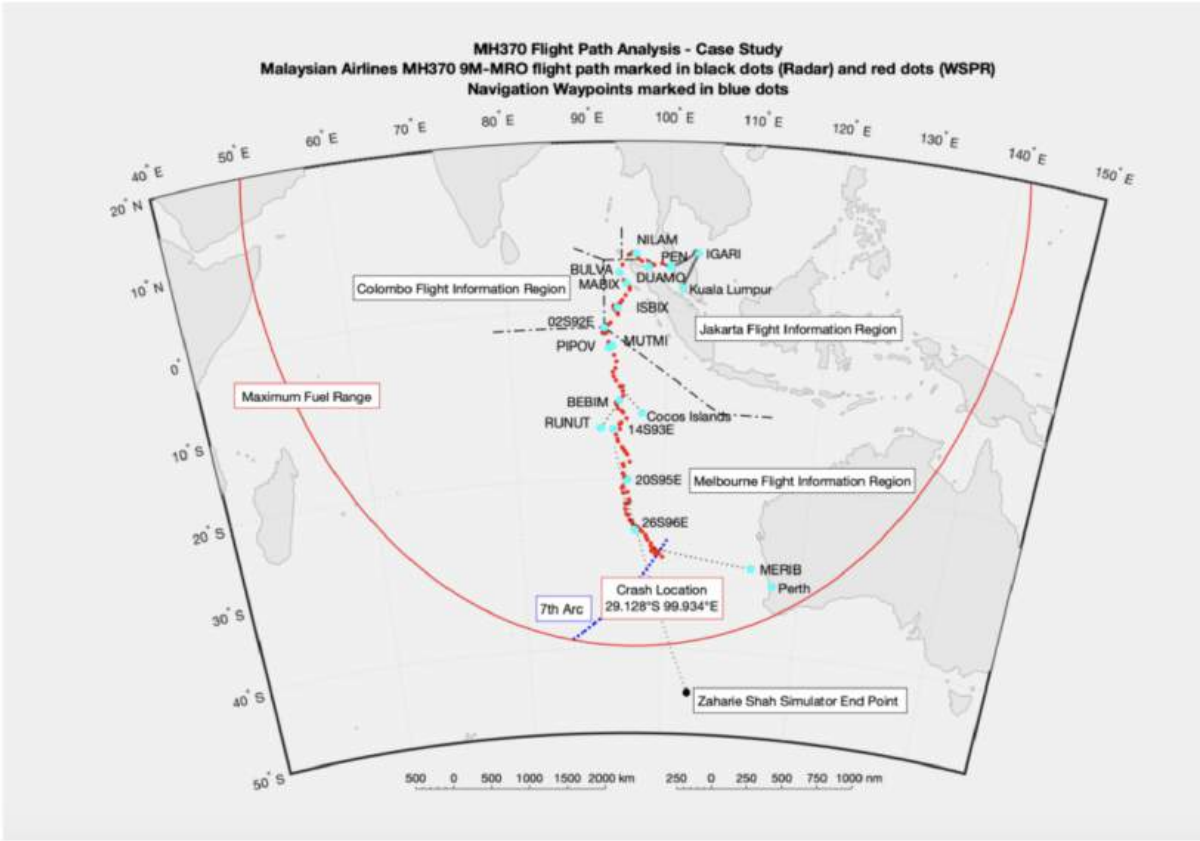
Nine minutes after this, HCM ATC attempted to contact the aircraft, only receiving mumbles in reply. Seven minutes after this, when the ACARS system was due to transmit a signal expected by ATC, nothing was received. This was when HCM ATC contacted KL ATC to ask if any verbal contact was established, which resulted in a reply of nothing. HCM ATC realised that the aircraft did not reach the waypoint of BITOD, which was supposed to be when it would contact HCM ATC. At around 01:52 local time, MH 370 reached the south of Penang Island, before turning towards the Strait of Malacca. Meanwhile, while KL ATC and HCM ATC were figuring out where the aircraft was, the Thai military were tracking it on radar. But miscommunication caused them to believe it was a regular commercial service being diverted back to Malaysia.

At 02:03 local time in Malaysia, the Malaysia Airlines Dispatch Centre sends multiple signals to the aircraft, to which no replies are received. A few seconds later, KL ATC asks the Malaysia Airlines Dispatch Centre if they have the location of MH 370, to which they state they have found it above Cambodia. At 02:15, KL ATC asks MADC if they have contact with MH 370, to which they reply that they can transmit signals and are receiving signals to and from the aircraft. 3 minutes later, KL ATC contacted ATC in Cambodia to ask them if MH 370 was in their control, but they said they do not have information or contact with the aircraft. It was later informed to KL ATC that the MADC could only predict the position of aircraft, not track it precisely. Since Cambodia was not on MH 370s planned route, the ATC got to work.

Above Penang Island, the cell-phone of First Officer Fariq Abdul Ahmed receives a signal, but no data could be found. KL ATC and HCM ATC contacted Chinese ATC between 03:30 and 04:25 local time, but they too did not have any information on the aircraft. At 02:25, a log-on request is sent from the aircraft to an Inmarsat satellite, followed by a “handshake” every hour, otherwise known as the transmission of data. At 05:09, Singapore ATC inquired about the aircraft as well, to see if they could help. At 06:32, after the aircraft missed its scheduled arrival in Beijing, the Kuala Lumpur Aeronautical Rescue Coordination Centre was activated to coordinate a rescue and to locate the aircraft. At 07:24, Malaysia Airlines released a press statement, exclaiming that MH 370 is missing. At 08:11, the last successful “handshake” was made with the plane. 8 minutes later, a log-on request is sent by the aircraft, before it loses power, and loses trace.

This is where most of the speculation arises. So many theories as to where the plane went after this resulted in a search area far too vast, which meant that, 4 years later, they still hadn’t found the plane. They called off the search after this. Many attempts were made to revive the search, but to no avail. Captain Zaharie had a flight simulator at home, in which it was found that he had flown this route in practice, but was deemed useless by authorities. In 2021, three people, by the names of Richard Godfrey, Dr. Hannes Coetzee, and Prof. Simon Maskell thought of using a technology called WSPR (Weak Signal Propagation Reporter) to find out the route of MH 370. This technology was developed by Joe Taylor, who received a Nobel Prize for Physics in 1993. He created this technology in 2008, and never believed it could be used for aircraft tracking. But, to his amazement, these three people have managed just that.

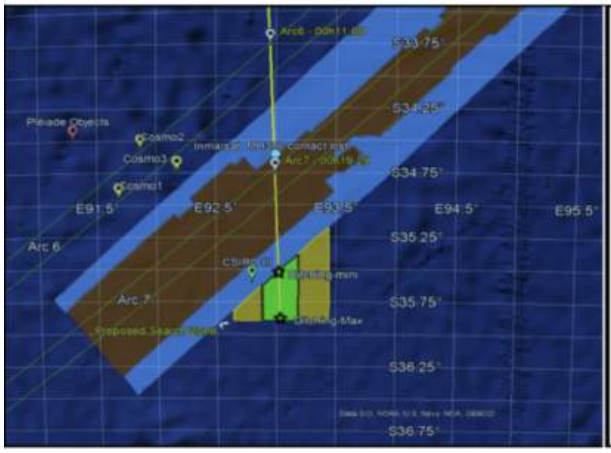
The way this technology works is by transmitting radio signals, but if an object comes in the way, there will be a small wave, or disturbance, in the signal. These three people used this technology, and worked on it for three years, before releasing a report and video in March of 2024 to resume the search for MH 370 in a specific area. Provided below is the detailed route taken by MH 370 and the area to search in.

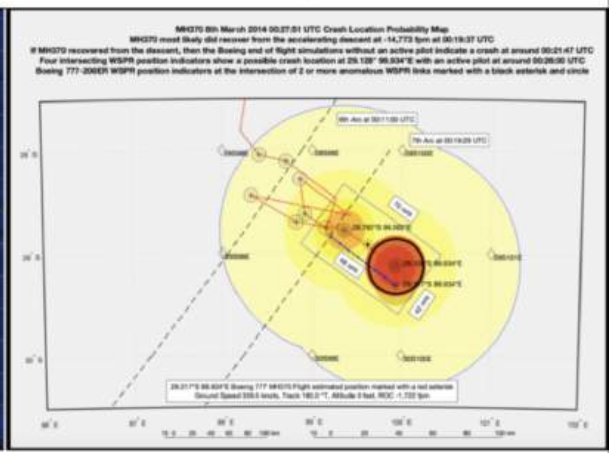


Joint Search Proposal - Petter Hörnfeldt - Mentour Pilot

Capt. Patrick Blelly and Jean-Luc Marchand
35.875°S 93.039°E - 1,200 km² - Green Area

Richard Godfrey, Dr. Hannes Coetzee and Prof. Simon Maskell
29.128°S 99.934°E - 2,800 km² - Red Area





Works Cited

Godfrey, Richard. “WSPR.” The Search for MH370, 5 April 2025, https://www.mh370search.com/category/wspr/. Accessed 10 April 2025.

Kelly, Megyn. Timeline of Malaysia Airlines Flight 370 - Wikipedia, https://en.wikipedia.org/wiki/Timeline_of_Malaysia_Airlines_Flight_370. Accessed 1 April 2025.

“Malaysia Airlines Flight 370 - Wikipedia.” Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Malaysia_Airlines_Flight_370. Accessed 20 May 2025.

“A NEW Trace! The FULL MH370 Story, so Far..” YouTube, 16 March 2024, https://www.youtube.com/watch?v=Y5K9HBiJpuk. Accessed 15 April 2025.

The Effect of Gender, Race, and Socioeconomic Biases on the UNOS Organ Transplant Waitlist and Organ Allocation By Joshua Devine

Abstract

The United Network for Organ Sharing (UNOS) is the main organ donation and distribution system in the United States. As of April 2025, there are 104,932 patients on the organ transplant waitlist (UNOS). The order in which patients fall on the waitlist is determined by more factors than just a first-come, first-served approach. It involves various medical tests and determinants in waitlist placement. However, despite these medical determinants, various disparities can be observed when it comes to one's wait time on the UNOS organ transplant waitlist. For example, when looking at gender, women face natural biological setbacks when it comes to tests that determine their placement on liver transplant waitlist. Women additionally often face setbacks while receiving a referral for organ transplantation, with physicians often influenced by social stereotypes of women and their health. In addition, there are also racial and socioeconomic biases that can be found in the UNOS waitlisting system, such as black patients being subjected to an average of one year longer on intermediate care in comparison to white patients in need of an organ transplant. Alluding to these findings, it is clear that the UNOS organ transplant waitlist system must be changed to eliminate all non-medical basis, to ensure that organs are being allocated to those who need them most, not who best fit the eyes of society.

Introduction

The United Network for Organ Sharing (UNOS) is the main organ donation and distribution system in the United States. UNOS utilizes a multi-tiered program for determining patient prioritization based on the national waiting list. As of April 2025, there are 104,932 patients on the organ transplant waitlist (HRSA). The order of the waitlist is determined by more factors than just a first-come, first-served approach. Specifically, blood type match, medical need, and tissue match are the primary factors in determining the list order. Although this approach may seem direct and unbiased, various factors may leak into the system to introduce bias from UNOS and healthcare providers. This discussion will examine how gender, race, and socioeconomic status influence one's placement on the UNOS organ transplant waiting list and one's likelihood of acquiring essential organs.

While the UNOS waitlist is only accessible by physician referral, unequal movement, mobility, and access to the waitlist still exist. The UNOS waitlist as of April 2025 included over 86% people in need of a kidney transplant, 9% needing a liver transplant, and 3% needing a heart transplant. Additional percentages consist of various other organs of need (UNOS). Although the racial composition of the UNOS waitlist compares similarity to national demographics (Organ Transplants and Black/African Americans), disparities in the rates of transplants conducted display underlying inequalities.

This review displays research and evidence exemplifying the effect of gender, race, and socioeconomic status on patients standing on the UNOS organ transplant waitlist and the chances of receiving a transplantation.

Results

Gender Inequality in Organ Transplantation

Based on survey data from the National Institute of Public Health, there is a large difference between the rates at which women donate organs, and the rates at which women need organ transplantation. Specifically women account for more than 60% of the organs that are donated, whether alive or deceased. However women only account for around 40% of those in need of organs (Yee et. al.). Women are very disadvantaged in the organ transplant process compared to men. In a study done by Amanda J. Vinson in association with Dalhousie University, after careful examination of the organ transplant waitlist over time, changes in position and rates that one received care displayed that women were less likely to be referred for transplant, particularly kidney transplant, less likely to progress on the waitlist in comparison to the rates at which men progressed, and less likely to receive the necessary treatment that they needed for survival (Vinson).

This inequality can be traced to the root of the UNOS waitlist: physicians. Physicians and primary care workers are the initial contact for patients needing an organ transplant. Patients must receive a referral from their physician to be placed on the United Network for Organ Sharing (UNOS) transplant waitlist. In a study conducted by BMC Nephrology, women on the kidney transplant waitlist were surveyed about their experiences with evaluation, referral, and waiting time on the UNOS waitlist. The study also used long-standing UNOS data as a comparison. Based on this survey, researchers were able to find that women encounter much greater doubt and skepticism when pairing a kidney transplant in comparison to men. Women had to jump through more hoops, such as additional appointments and referrals, and undergo a much more extended process to achieve a referral compared to male patients (Gompers et al.). As seen in the same study, women face much lower rates of referral, evaluation, and waitlisting for kidneys. Once waitlist status was achieved, women waited for much longer periods in comparison to men, displaying gender disparities in movement and acceptance to the organ transplant waitlist (Gompers et al.).

These biases are not only the result of flawed medical analysis and testing but can also be seen in physician biases and external societal influences. A previous study conducted by Amanda J. Vinson, in collaboration with Dalhousie University, also examined how physicians perceive their patients during care and referral. Based on female patient testimony and waitlist standings, she was able to draw conclusions about the experience that many women faced during the transplantation process. Vinson was able to observe that healthcare providers during primary screening often used a patient's perceived frailty as a factor in waitlist referral, and this factor had a significantly greater impact on women, contributing to the discrepancy in referral rates between women and men, especially among the elderly. Vinson defined frailty in her study as an

age-related state of increased vulnerability to adverse health outcomes (Gordon & Hubbard). Based on her findings, she found that older women tended to be rated as more frail than men of the same age, purely based on social stereotypes and surface-level views of their health.

These ratings and decisions by physicians inherently set these women further and further back on the waitlist, as the frail state the doctors saw them in could pose a risk for post-transition complications (Vinson). This view of women's concerns being seen as lesser is a view rooted deep in our society, and one that must be pushed out of the minds of healthcare providers when determining status and conditions used in waitlist determination.

These disparities persist beyond the initial referral period for women. In terms of liver transplantation, waitlist position and urgency are determined by a Model for End-Stage Liver Disease score (MELD score). The MELD score ranges from 6, indicating less severe symptoms and lower urgency, to 40, meaning the patient is in dire need of a liver transplant. The MELD score is determined by the patient's levels of bilirubin, albumin, sodium, and creatinine (UNOS). While levels of bilirubin, albumin, and sodium in the body are typically the same when comparing men with women, creatinine levels (determined by muscle mass) are naturally higher in men, as they biologically possess a greater amount of muscle mass (Rogers et al.).

In a study conducted by the Routledge, Taylor, and Francis group in New York and London, comparing MELD score levels in men and women based on these four substance levels in the body, a systemic bias can be observed in the liver transplant system based on the MELD score. Since impacted livers are usually unable to cleanse creatinine from the blood at healthy rates effectively, men have greater creatinine levels during liver impairment. This physical and biological difference contributes directly to a greater MELD score in a man, even with the same liver disease severity as a woman, due to their increased creatinine production even in a healthy body (Rogers et al.). This UNOS-required and implemented system leads to more seriously affected women being pushed lower and lower on the organ transplant list by less severely affected men who have naturally higher MELD scores due to naturally higher creatinine production levels.

Racial Disparities in Organ Transplantation

Despite the race of a donor not being a contributing factor in organ survival or viability rates, and matching the race of the donor and recipient not being a medical necessity (Middleton), some racial populations biologically face higher rates of specific diseases that can contribute to increased need for organ transplants. Based on specific data collected by the US Department of Health, Black patients are 1.4 times more likely to be diagnosed with diabetes in comparison to white non-hispanic patients (Office of Minority Health). This increased likelihood contributes directly to increased chances that these patients will experience kidney disease, requiring kidney transplantation (Wesselman et al.). Despite this higher need for care, black patients have a lower percentage of being selected as kidney transplant recipients.

According to a research study at the University of Pittsburgh Medical Center (UPMC), this disparity is well-demonstrated. Looking at patients at the UMPC, researchers were able to

find that of the white patients awaiting kidney transplants, 38% received the transplant that they needed. Looking at black patients waiting for transplants, only 24% received a transplant. The difference in white patients having a 14% greater chance of receiving a transplant displays the underlying biases and racial differences in the organ transplant system and waitlist (Wesselman et al.).

In the same study conducted at the UMPC, further evidence of this disparity of organ transplant rates due to race can be seen. Examining patients waiting for transplants and the care they received during this period revealed variations in care. White patients, on average, spent much less time on dialysis, a form of kidney replacement used in patients with severe kidney disease awaiting a transplant, compared to black patients. Black patients were subjected to this intermediate care for much more extended periods than patients of other races, putting their health at risk as they waited for more extended periods for a transplant. Based on data from the UMPC, this difference in the period of dialysis instead of receiving a transplant is apparent. Black patients, on average, spent approximately one year longer on dialysis than did white patients, with their risk of death from kidney disease increasing every day. This difference in wait time of over a year for patients looking to receive a kidney transplant directly displays the bias in the waitlist system.

Socioeconomic Disparities in the Organ Transplant System

When taking factors such as race and gender out of the equation, socioeconomic status, tracked using markers such as wealth and geographic location, is a significant determinant in one's likelihood of receiving life-saving organs. One's insurance plan's coverage, financial status, post-operation costs, and overall accessibility and proximity to transplant centers all become significant factors in a patient's ability to access transplant procedures.

Financial hurdles account for a large amount of obstacles when it comes to patients receiving an organ transplant (Hays et al.). According to a research study on Organ Transplant list analytics by the National Library of Medicine, considerable ethical issues arise from the fact that financial considerations and a patient's ability to pay influence how an organ is allocated to one patient over another.

The push for financial neutrality is strong in the medical community, allowing medical need to be the sole factor in organ allocation, rather than financial ability. The issue of those with more money getting priority for care is growing ever more prevalent in the modern world. A study done by Hod and Goldfarb saw similar issues to this. They analyzed groups of various socioeconomic status and their outcomes, including those who obtained and live with organ transplants. Through their research, they found that low income individuals faced improper care and acute organ rejection at a rate of 36%, compared to high-income patients who received a rejection rate of 27%. (Hod & Goldfarb). Individuals with a higher income have nearly a 9% greater chance of surviving the process of organ transplantation, directly highlighting the disparity in the organ transplant process and care given between those of different socioeconomic status.

Geographic factors additionally play a critical role in this socioeconomic disparity. Organ transplant centers often operate out of large hospitals, primarily located in cities and urban areas. Those living in more rural areas tend to have lower incomes in comparison to those living in urban centers, limiting not only their proximity to medical centers, but their ability to reach them at all. This poses a challenge for rural residents needing a transplant, sometimes living hours away from the closest transplant center. Where an organ transplantation patient is listed to receive their procedure is noted by UNOS to be one of the most significant factors when determining organ allocation, giving way to large disparities within the system. Due to the organ's time of viability, often not more than a couple of hours, UNOS must work quickly to make sure that an organ is allocated enough time for a successful transplant. However, with these quick decisions, rural residents are often overlooked as candidates, seemingly too far away to safely transplant the organ to (Hackmann et al.). This constant overlooking of rural residents can lead to years on the waitlist, with individuals being automatically dismissed from consideration time and again due to their distance from a transplant hospital.

One's socioeconomic status has a significant and influential effect on their access to organ transplantation. Financial barriers and assessments based on income prevent those of lower socioeconomic status from even having the chance of an organ transplant. Additionally, one's proximity to a transplant center is listed by UNOS itself as a factor, a factor that has led to patients being left out of consideration for transplant due to where they live. Socioeconomic status should not be used as a determinant for organ allocation, which should focus solely on medical need.

Conclusion

The Organ Transplant System is a pivotal and essential aspect of global medical care. Without it, people would never be able to receive organ transplants and the care they need for survival. Kidney and liver disease would become deadly in UNOS’s absence, and heart issues would become far more fatal. There is no doubt that the Organ Transplant System is essential, operating to serve the people and their needs. However, the biases and factors considered for organ allocation raise questions about how it should be determined.

Considering the determining factors, purely medical-based systems must be implemented to mitigate some of these biases. Making sure that one's physical appearance or assumptions based on age or looks are excluded from the referral and scoring process of organ necessity, are steps in the right direction to unbiased organ allocation.

Even in some of the medical scales themselves, disparities can be found. When evaluating MELD scores for liver disease, it is essential to standardize all factors in these scores across men and women, ensuring that biological components such as muscle mass or nutrient production rates do not influence the assessment of how a disease affects the body. These scales for organ urgency must purely reflect the urgency in question, not the patient's sex or their natural reproductive sex of the patient or their natural production systems. Movement must be

made to eliminate these weighted determinants in the allocation system of organs to cleanse the organ transplant system of further disparities.

A patient's race must also not play a factor in movement on the UNOS waitlist. The fact that black patients often wait up to a year longer on the waitlist than non-Hispanic white patients directly shows the action that must be taken to remove race as a determinant in the organ selection process. The stagnation in movement on the waitlist of patients of color displays long-lasting disparities that must be eliminated for an ethical waitlist to be achieved. Changes such as a race blind evaluation system and removing that factor from all documents that it is not needed in, can help to eliminate these issues of racial biases of the organ transplant waitlist.

One's low financial income can be seen to greatly reduce the likelihood of having an organ transplant and moving up on the waitlist. This factor turns the medical field against those with lesser income and makes them much less likely to actively or even pursue the case in which they need it. In addition, work must be done by UNOS, insurance agencies, and society itself to reduce the factor of distance and proximity in organ allocation. Funding for travel and heightened alert time of organ availability could all help to remove this factor as much as possible, working with organs short viability. Work must be done so that the determinant of distance from a medical center can be reduced to only come into play when the organ's viability is at risk, and organs must not be given to those most convenient, but to those in most need.

Works Cited

Gompers, A., Rossi, A., & Harding, J. L. (2024). Intersectional race and gender disparities in kidney transplant access in the United States: a scoping review. BMC Nephrology, 25(1). https://doi.org/10.1186/s12882-023-03453-2

Gordon, E. H., & Hubbard, R. E. (2019). Differences in frailty in older men and women. Medical Journal of Australia, 212(4), 183–188. https://doi.org/10.5694/mja2.50466

Hackmann, M., English, R. A., & Kizer, K. W. (2022, February 25). Saving More Lives and Enhancing Equity with Deceased Donor Organ Allocation Policies. Nih.gov; National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK580021/

Hays, R., Rodrigue, J. R., Cohen, D., Danovitch, G., Matas, A., Schold, J., & LaPointe Rudow, D. (2016). Financial Neutrality for Living Organ Donors: Reasoning, Rationale, Definitions, and Implementation Strategies. American Journal of Transplantation, 16(7), 1973–1981. https://doi.org/10.1111/ajt.13813

Hod, T., & Goldfarb-Rumyantzev, A. S. (2014). The role of disparities and socioeconomic factors in access to kidney transplantation and its outcome. Renal Failure, 36(8), 1193–1199. https://doi.org/10.3109/0886022X.2014.934179

HRSA. (2025, March). Organ donation statistics. Www.organdonor.gov; Health Resources & Services Administration. https://www.organdonor.gov/learn/organ-donation-statistics

Middleton, J. (2023, August 25). Does Race Matter in Organ Transplants. Donor Alliance. https://www.donoralliance.org/newsroom/donation-essentials/multicultural-groups-donation-does-race-matter-in-organ-transplants/

Office of Minority Health. (2024). Diabetes and Black/African Americans. Office of Minority Health. https://minorityhealth.hhs.gov/diabetes-and-blackafrican-americans

Organ Transplants and Black/African Americans. (2024). Office of Minority Health. https://minorityhealth.hhs.gov/organ-transplants-and-blackafrican-americans

Rogers, W., Carter, S. M., Jackie Leach Scully, Entwistle, V. A., & Mills, C. (2023). The Routledge Handbook of Feminist Bioethics. Routledge.

Vinson, A. J. (2022). Gender Disparities in Access to Kidney Transplant: Inequities in the Inequity. Kidney International Reports, 7(6), 1145–1148. https://doi.org/10.1016/j.ekir.2022.03.034

Wesselman, H., Ford, C. G., Leyva, Y., Li, X., Chang, C.-C. H., Dew, M. A., Kendall, K., Croswell, E., Pleis, J. R., Ng, Y. H., Unruh, M. L., Shapiro, R., & Myaskovsky, L. (2021). Social Determinants of Health and Race Disparities in Kidney Transplant. Clinical Journal of the American Society of Nephrology, 16(2), 262–274. https://doi.org/10.2215/cjn.04860420

Yee, E., Seyedeh Maryam Hosseini, Duarte, B., Knapp, S. M., Carnes, M., Young, B. A., Sweitzer, N. K., & Breathett, K. (2021). Sex Disparities in Organ Donation: Finding an Equitable Donor Pool. Journal of the American Heart Association, 10(19). https://doi.org/10.1161/jaha.121.020820

Analyzing the Effect of Match Pressure on Serve Placement in Professional Men's Tennis: A Composite Pressure Index Approach By Justin Tran

Abstract

Serve placement is one of the most critical tactical decisions in professional tennis, but individual point-level psychological pressure has not been studied. This paper presents a new measure of the Composite Pressure Index (CPI) that incorporates baseline chances of winning the game along with other important situational variables including break-points, set-points, match-points as well as tiebreak points. Based on a dataset of 423,728 point-level observations of ATP Tour events (2020-2024) we illustrate that the CPI has a highly significant correlation with first-serve success ($r = -0.150$, $p < 0.001$) and rate of double-faults ($r = +0.178$, $p < 0.001$). Under high CPI, players were found to manipulate serve direction and reduce T serves by 3.2 % points and increase wide serves by 3.0 % points. These movements were especially noticeable on the clay courts. Low-ranked players (ATP 501-2000) had stronger strategic aberrations, indicating that experience acts as a mediator of psychological stress reactions. This study provides empirical support for psychological theories of performance under pressure and offers a novel metric that can be applied to player preparation and tactical analysis.

Keywords

Tennis Analytics, Psychological Pressure, Serve Placement, Decision-making, ATP Tour, Composite Pressure Index (CPI)

Introduction

The serve placement strategy on the tennis court is a critical strategic move that has a direct effect on the point building and the match patterns especially in the professional scenario. The serve angle, e.g. down the T, wide, or body, can have a devastating effect on the capabilities of the receiver to make a decent return of the ball, and therefore it will affect the length of the rally, type of shots to be used, and point dynamics. Service placements can strategically control the flow of a match with players constantly changing their serve placement depending on situational factors. Past studies grasp an overview of biomechanical factors, including serve speed, spin, and precision (Elliott et al.), but they do not take into consideration the psychological and strategic aspects of serve choice when under pressure. Specifically, the mental part of serving particularly in key points has not been well researched. In professional tennis, high-pressure states, e.g. break-points, set-points, or match-points are crucial stages; how well one serves and what changes to use on serves may become a decisive factor in the result of the game. The psychological power of the player to change his serving game in these circumstances is usually what separates the professionals and the novice player.

In the traditional studies of match pressure, researchers usually take a binary division including break-point, tiebreak, or match-point status, to determine pressure moments (Peña López, J.). Although these divisions are helpful, they do not address the fact that the level and

situational stress of psychology differs as a match progresses. Tennis pressure is multi-dimensional and its effect cannot be defined narrowly to only high stakes moments but is also influenced by accumulating factors, such as the score, match progress, the surface, and momentum. The mental effect of these stresses can greatly change the process of the decision, especially when a player intends to serve the ball and where to serve it. In light of these nuances, a more dynamic and continuous measure of psychological pressure, applicable within the whole of a match, is becoming increasingly necessary as an approach that is superior to considering only stationary, binary measures.

In order to fill this gap, the current study offers the Composite Pressure Index (CPI), a new measure that quantifies psychological pressure at the individual point level by combining probabilistic models with contextual modifiers. CPI has the goal of providing a more detailed and moment-by-moment estimate of pressure, using game score, break-points, set points, match points, and tiebreak score. The cumulative effect of pressure on various match contexts enables CPI to observe the development of pressure throughout a match and time-varyingly affect the serve placement strategy of a player.

The main purpose of this study is to investigate the correlation between CPI and pressure-sensitive serve results, specifically first-serve success and double-fault percentages. This study also examines how players change their serve position, predictably, as they experience psychological pressure. As a case in point, do the players tend to play safer by widening the serve or serving to the body under pressure as it has been shown in previous studies (Cross & Pollard)? This study aims to give a contribution to this rapidly developing area of research relating to the analytics of tennis performance and provides further methodological instruments utilized by coaches, analysts, as well as sport psychologists who can apply the CPI to gain more insight into the mechanisms of how psychological pressure affects serve strategy. Recent innovations in tennis statistics have made it clear that performance under pressure can be crucial to the outcome of a match. As an example, a recent study by (Mesagno et al.) refers to the importance of psychological pressure in tennis, analyzing the implications of cognitive and emotional reactions to tennis performance. The results of their study indicate that the psychological conditions of the players in critical situations of the game have a considerable impact on their decision-making, including the placement of a serve (Mesagno et al.). Along the same lines, (Cowden et al.) also examine the correlation between mental toughness and match outcomes, concluding that more mentally resilient players tend to change their strategies under pressure and switch up their serves in tense points (Cowden et al.).

Moreover, new computational models have emerged to gain better insight into the interaction of psychological variables with performance outcomes. (Harris et al.) applied game-theoretic models to measure the impact of psychologically pressurized situations on player performance, and they showed that players tend to adjust the behavior when faced with a high pressure situation, such as adjusting serve locations (Harris et al.). These latest findings reinforce the significance of continuous and subtle pressure measures in elite-level tennis play.

The proposed research contributes to this line of research by presenting the CPI, a new instrument that provides a continuous and meaningful method to measure psychological pressure in professional tennis. This way, it expands our knowledge on the impact of pressure on the tactical decision-making process of professional players especially serving positions and serves as a rich source of information in preparing players, mental training, and post-match analysis. The results of this study will contribute to reducing a major gap in the literature by providing a direct association between particular tactical choices, like serve positioning, and psychological pressure and provide coaches and analysts with usable recommendations related to improving psychological pressure conditions in support of expected player behavior.

Methods

Source and preparation of data

In this paper, we utilize publicly available data from Jeff Sackmann on tennis_atp on GitHub (https://github.com/JeffSackmann/tennis_atp) that provides detailed point and match-level metadata on all ATP Tour events. Specific datasets used in the analysis were the following ones:

- charting-m-points-2020s.csv: In this file the information is provided point-by-point: direction of the serve, score of the game e.g., 40-15, 30-30, score of the sets e.g., 6-4, 4-4 and score of the match that is the number of sets won by each player.
- charting-m-matches.csv This file contains the metadata of each match with respect to the surface type (clay, grass, or hard), round, and date of match.
- atp_players.csv: Identification data of participants, i.e., player IDs and full names.
- atp_rankings_current.csv: ATP player rankings at time of the matches.

Pandas library of Python was used to process the datasets. The tables of point-level and match-level were joined by match_id, and the names of the players were converted into their normalized form, with the underscore replaced by the space character. A two-stage join was used to match the names of the players to their respective playerIDs and ranking. In the analysis, second serves were also analyzed where there was a missing first-serve code so that the serve direction data would not be discarded. There were a total of 423,728 point-level observations used based on ATP Tour events (2020-2024).

Feature Engineering

Serve directions were standardized with 3 categories: wide, body, and T (4 = wide, 5 = body, 6 = T). About 2 % of rows with an unknown serve direction code were discarded. Contextual flags were formulated systematically through the match score data to find high pressure situations. To be more precise, the following flags were designed:

1. Break point: A flag for points occurring at break-point circumstances.

2. Set point: A flag for points occurring in set-point situations.
3. Match point: A point flag that represents points in a match-point situation.
4. Tiebreak: Points flags used in tiebreak play, either in the first half of the tiebreak (points < 5) or in the second half of the tiebreak (points 5 or more).

The matches were additionally split by the type of the surface (clay, grass, hard) and then the players were organized into some ranking categories (Top 100, 101-500, 501-2000) according to the data in atp_rankings_current.csv.

Construction of Composite Pressure Index (CPI)

The Composite Pressure Index or CPI was made by combining baseline game-winning probabilities with certain contextual modifiers. Baseline (B) component was estimated based on the game-winning probabilities outlined by Klaassen & Magnus. To represent the influence of psychological pressure at the point level, contextual modifiers, which represent different pressure situations, were given a certain weight. These modifiers are:

1. **Break point**: A weight of +0.20 was assigned if the point occurred at a break-point.
2. **Set point**: A weight of +0.20 was assigned if the point occurred at a set-point.
3. **Match point**: A weight of +0.50 was assigned if the point occurred at a match-point. **Tiebreak**: Early tiebreaks (points < 5) were assigned +0.08, and late tiebreaks (points ≥ 5) were assigned +0.28.

Moreover, a momentum factor (M) that measured set progression was also added to the CPI formula with values varying between +0.10 to -0.10 depending on the set deficit or progression. To match the overall importance of the match according to the CPI a scaling factor (S) was applied. These contextual factors have been optimized using a grid-search approach on a 60-40 train-test data split in order to maximize the correlation between the CPI and observed pressure-sensitive variables like first-serve success and a double-fault rate.

$$CPI = B + (W_{BP} \cdot I_{BP} + W_{SP} \cdot I_{SP} + W_{MP} \cdot I_{MP} + W_{TB} \cdot I_{TB} + M) \times S$$

Where

B = Baseline pressure from Klaassen & Magnus probabilities

$W_{BP}, W_{SP}, W_{MP}, W_{TB}$ = Weights (0.20, 0.20, 0.50, 0.08/0.28 respectively)

$I_{BP}, I_{SP}, I_{MP}, I_{TB}$ = Binary indicators for break-point, set-point, match-point, and tiebreak context, respectively

M = Momentum factor (±0.10 based on set deficit/progression)

S = Scaling factor adjusted for match context

Parameter Optimization

Grid-search optimization was used to optimize the CPI parameters and maximize the correlations with important server outcomes. In particular, the weights of each CPI modifier were tuned according to a split of train-test data at a 60-40 ratio. The optimization procedure was expected to reach as high as possible the level of correlation with two important measures that are related to the serve, which are first-serve success and double-fault rates. These correlations were maximized empirically by establishing the weights of CPI modifiers.

Statistical Analysis and Outcome Metrics

First-serve success rate and the rate of the double-fault are the main outcome measures of this research study, which were correlated with the mean CPI of each match. These relationships were evaluated using Pearson correlation coefficient, and bootstrap confidence intervals were computed (5000 iterations) in order to give us strong estimates of correlation coefficients.
In addition to the correlation analysis, proportions of each serve-direction were calculated across various CPI terciles (low, mid, high), and statistical differences in serve direction between terciles assessed with chi-square tests. This discussion was able to give an analysis on the influence of psychological pressure on strategic serving locations with high pressure and low pressure.

Serve-Direction Analysis

The serve direction proportions were compared between the CPI terciles (low, mid, high) to determine whether there is any significant change in tactical choice. This analysis was conducted on the basis of surface type and ranking of players, and focused on the psychological pressure on players according to the nature of the match. Serve direction data were also analyzed to identify any significant differences in behavior in higher- and lower-ranked players, especially the ways the two categories of players could modify their serve strategy in pressure situations.

Results and Discussions

The Composite Pressure Index (CPI) distribution revealed a right-skewed profile, consistent with expectations that low-pressure points are more frequent in professional match play. The mean CPI level was 0.56 (SD = 0.42), which shows that the majority of the points were played under moderate pressure, whereas very high-stress situations are not so common as shown in Figure 1.

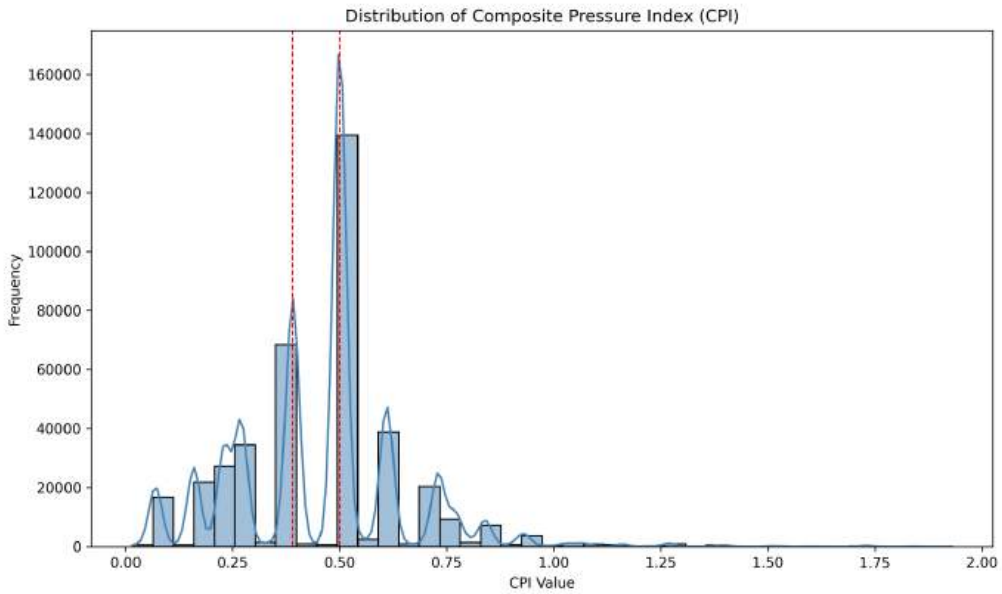


Figure 1: Histogram of CPI distribution

A validation test proved that CPI was closely related to recognized metrics of pressure serves. First-serve success was found to exhibit an intermediate negative correlation with CPI ($r = -0.197$, 95% CI [-0.220, -0.173], $p < 0.001$) as shown in Figure 2 and the rate of double-faults had a positive correlation ($r = +0.202$, 95% CI [+0.178, +0.227], $p < 0.001$) shown in Figure 3. These patterns affirm the CPI's construct validity as a psychological pressure indicator.

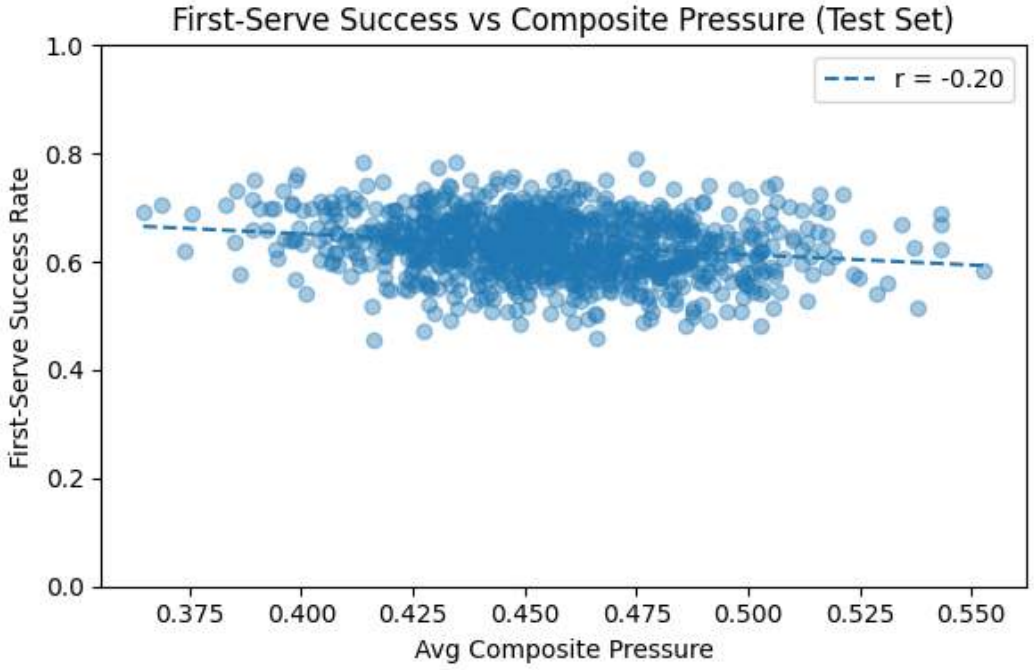


Figure 2: First Serve success rate vs Composite Pressure

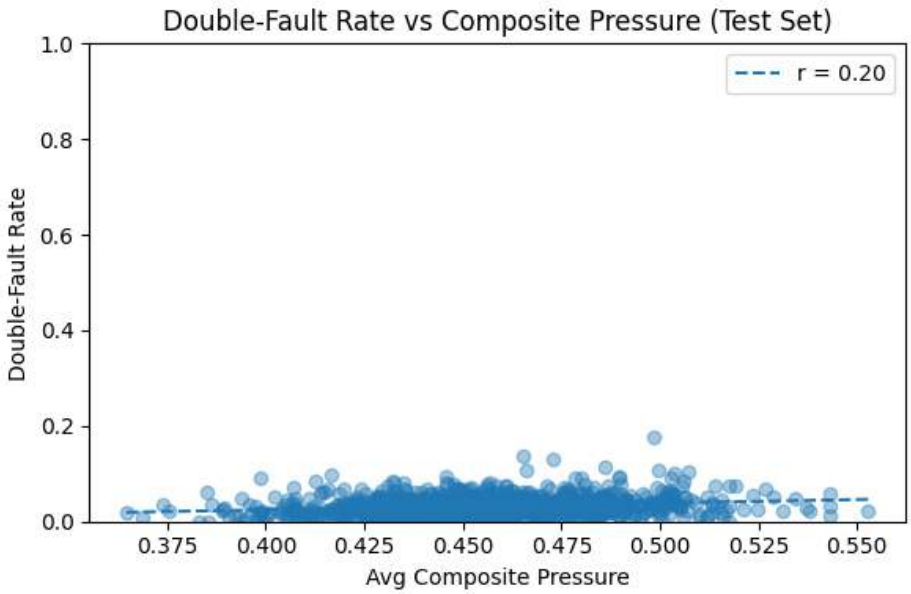


Figure 2: Double Fault rate vs Composite Pressure

Analyzing serve-directional preferences by CPI tercile as shown in Figure 3 and 4, revealed that players in high-pressure environments (top third), showed a significant change in their tactics; they decreased T serves by 3.2 % points and increased wide serves by 3.0 % points. These changes were found statistically significant by chi-square analysis ($\chi^2(4) = 512.6$, $p < 0.001$).

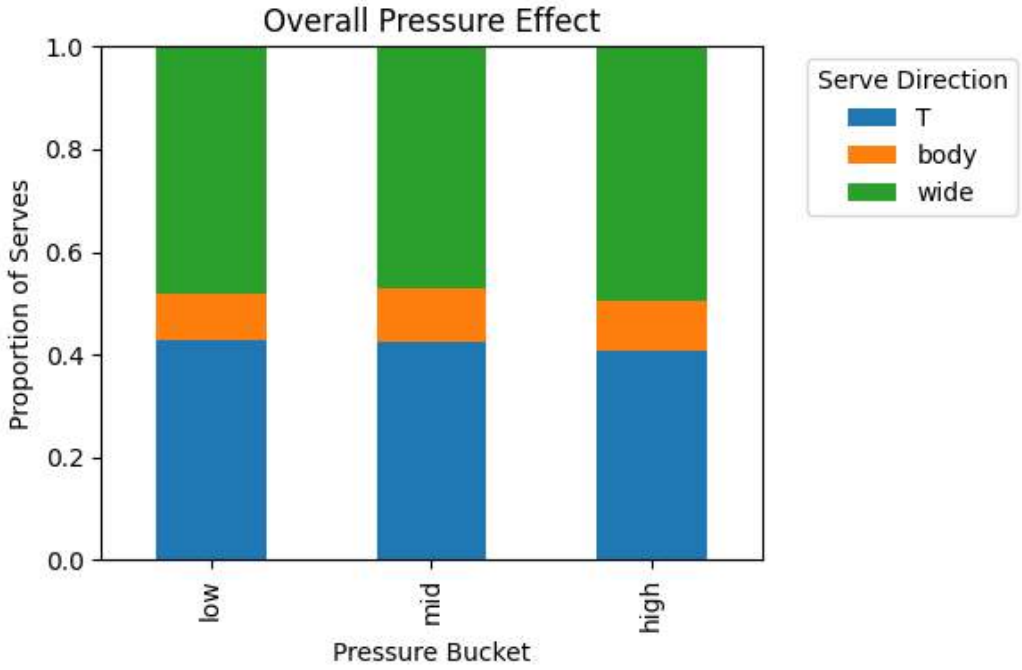


Figure 4: Bar chart of serve direction proportions by CPI tercile, highlighting shifts from T to wide serves under pressure.

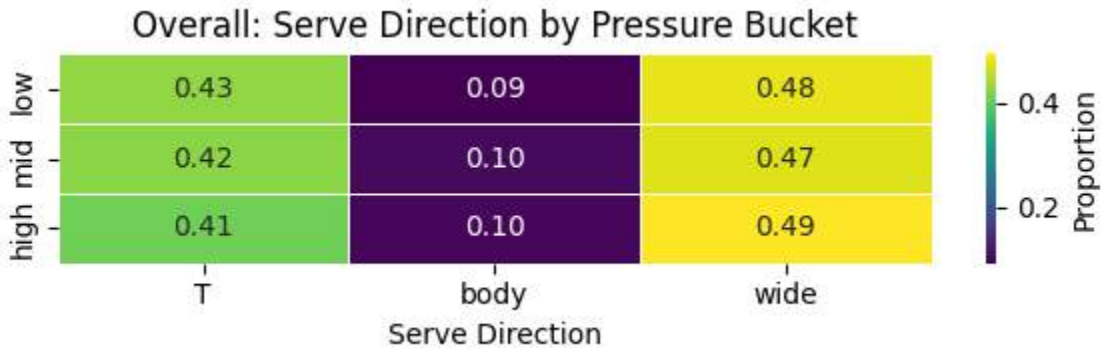


Figure 5: Heatmap of Serve direction by Pressure bucket

A surface-specific analysis indicated that such strategic adjustments were amplified by clay courts. When CPI was high, T on clay declined by 4.5 % points, and wide T on clay increased by 4.0 % points, but the effects on grass and hard courts were small but significant as indicated in Figure 6. This indicates surface responses through pressure, which is probably due to varied serve-return forces.

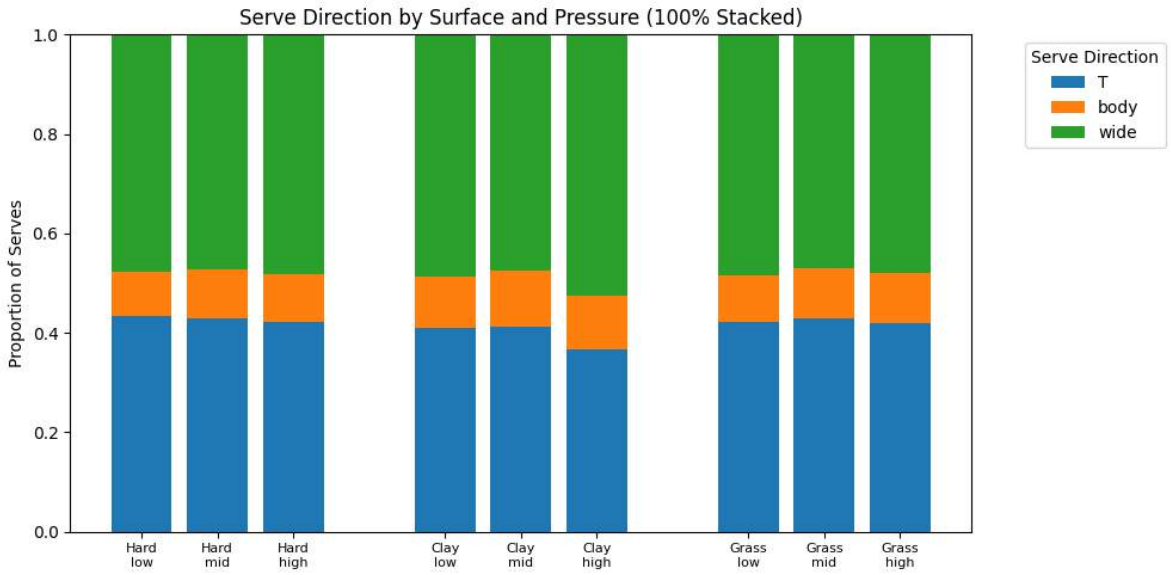


Figure 6: Serve Direction by surface and pressure

The player ranking tier further moderated the pressure–strategy relationship. The direction changes between high CPI and low CPI were most significant in lower-ranked players (ATP 501-2000), which indicates high pressure vulnerability to pressure-provoked tactical variation. Top-50 players, in turn, had more consistent serve tendencies, which may be a sign of resilience gained with experience.

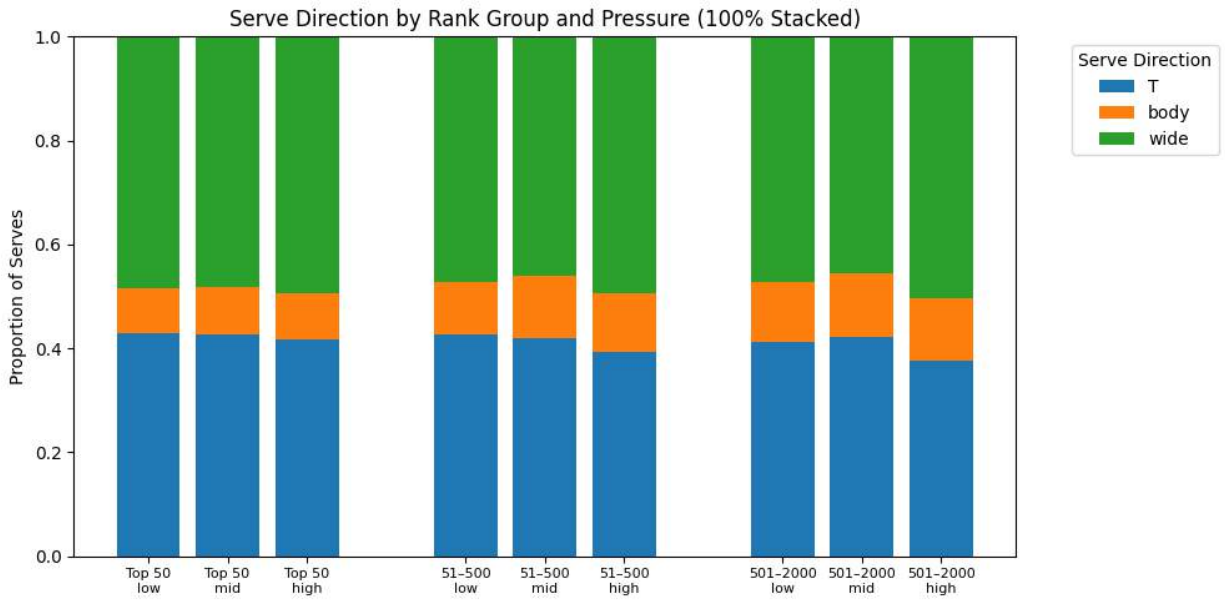


Figure 7: Serve direction by group ranking and pressure

In this study, the Composite Pressure Index (CPI) is the refined, contiguous measure of psychological pressure at the point level, significantly advancing the binary measures of the past, which relied on indicators like break-point or tie break status. Through a combination of baseline win probabilities and situational modifiers, CPI encompasses the dynamic multi-dimensionality of in-match pressure. Its construct validity in terms of the supposed reliability of this performance-pressure indicator is locking into the significant negative correlation with first-serve success and the positive correlation with the rate of a double-fault. Serve-direction analysis showed a sequential modification in the tactical sense with high CPI because players moved beyond T serves to wider positions. This defensive, unaggressive approach is probably aimed at: maximising serve margin, lengthening the returner to an angle, and also reducing direct-return winners. Clay courts had the strongest impact as ball pace is slower and the bounce is higher reducing the penetration of T serves and boosting the efficiency of wide serves to interfere with the opponent positioning.

The ranking-tier investigation revealed that, tactically and under pressure, less-rated players (ATP 501-2000) presented more tactical deviations as compared to the top-50 players, who were more strategically constant. This result implies that competition experience and mental resilience are important in maintaining stress-sensitive tactical patterns. These findings are in line with the current literature on mental toughness which shows that top professional players are more likely to engage in pre-made routines and possibly serve routines to resist bias-forced cognitive intrusion. All of these findings further support the validity of CPI as not only a means by which pressure should be ascertained but as a model of how players create a change in tactics to suit the psychological demands, playing field requirements, and game position.

Conclusion and Future Work

This paper presents the Composite Pressure Index (CPI) as a consistent and interpretable method of measuring match pressure in professional tennis, providing a continuous, context-sensitive index that will reliably measure the impact that pressure has given to serve placement. By revealing surface-specific and experience-related differences in tactical response, CPI provides immediate value for coaches, analysts, and sport psychologists in designing targeted, pressure-specific training and performance evaluations. Future studies can generalize CPI by including serve-speed, spin, and positioning of the returning player data to reflect three-dimensional tactical change; use the metric in WTA, doubles, and junior tournaments to test wider applicability; and incorporate dynamically estimated ranking models to reflect in-bound form and tournament response to stress. The predictive capacity of CPI could further be validated and causal relationships between pressure, decision-making and execution clarified as experimental research attempts to replicate CPI-derived pressure situations in controlled conditions.

Works Cited

Elliott, Bruce, Machar Reid, and David Whiteside. "Biomechanics of groundstrokes and volleys." Tennis medicine: A complete guide to evaluation, treatment, and rehabilitation. Cham: Springer International Publishing, 2019. 17-42.

Peña López, Javier. "Analysis of the service as a performance factor in high-level volleyball and beach volleyball." (2013).

Cross, Rod, and Graham Pollard. "Grand Slam men's singles tennis 1991-2009 serve speeds and other related data." Coaching & Sport Science Review 16.49 (2009): 8-10.

Mesagno, Christopher, Adwoah A. Hammond, and Matthew A. Goodyear. "An initial investigation into the mental health difficulties in athletes who experience choking under pressure." Psychology of Sport and Exercise 74 (2024): 102663.

Cowden, Richard G. "Competitive performance correlates of mental toughness in tennis: A preliminary analysis." Perceptual and motor skills 123.1 (2016): 341-360.

Harris, David J., et al. "Psychological pressure and compounded errors during elite-level tennis." Psychology of sport and exercise 56 (2021): 101987.

Klaassen, Franc, and Jan R. Magnus. Analyzing Wimbledon: The power of statistics. Oxford University Press, 2014.

Sackmann, Jeff. tennis_atp: ATP Tennis Rankings, Results, and Stats. GitHub, https://github.com/JeffSackmann/tennis_atp. Accessed 10 Aug. 2025.

Genetic Engineering and Bioethics: Balancing Progress with Ethical Limitations

By Chris Sullivan

Authors Note

This article was completed as part of a research initiative conducted through the Righetti Engineering Organization with the guidance of Professor of Physics and Stanford alumni, Colin Rodriguez, who serves as the program's faculty advisor. The project is supported by the collaborative efforts of the Righetti Engineering organization, exemplifying the program's commitment to cultivating student-driven research and providing hands-on experience in engineering, which serves to support future career opportunities and further academic pursuits in the field. Special thanks are extended to the Ernest Righetti High School administration for their ongoing support, and to other leaders of Righetti Engineering for their valuable guidance and contributions throughout the study. Please address correspondence to: Christopher Sullivan, Righetti Engineering Organization, Ernest Righetti High School, 914 East Foster Road, Santa Maria, CA 93455. Email: chrissullivan0905@gmail.com.

Abstract

The idea of widespread genetic engineering in society has tremendous upside, as it offers advancements in treating diseases and enhancing human traits such as intelligence, physical appearance, and strength. However, the ethical and social implications of genetically modifying human DNA for disparate purposes are highly contentious. This report investigates the ethical and moral concerns surrounding genetic modifications, particularly in the context of enhancing intelligence and physical attributes. It attests that while genetic engineering could offer significant medical benefits, its use for enhancement purposes may exacerbate social inequalities, undermine fairness, and threaten personal autonomy. The study also covers the possibility of a genetically advantaged class forming, the risks of reinforcing harmful societal standards, and the devaluation of human achievement. It concludes that genetic engineering should be strictly limited to medical applications by law to avoid societal harm and maintain fairness and equity.

Genetic Engineering Advances and Mechanisms

In the summer of 2019, a transformative jump in science and human progress unfolded, as Victoria Gray, a thirty-nine-year-old woman, had over two billion newly edited CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) stem cells infused into her body via a catheter in her chest ("An Answered Prayer: Warrior Reflects on Her Experience as the First Patient Treated with CRISPR for Sickle Cell Disease"). Gray had been suffering from Sickle Cell disease—a genetic disorder causing red blood cells to form with an abnormal shape, leading to blockages in blood flow and potential organ damage—since a young age, experiencing symptoms such as chronic pain, fatigue, infections, and episodes of extreme pain known as sickle cell crises. Because of the operation, doctors were able to prompt the production of

healthy red blood cells and realize prominent results: “The disease that had plagued Gray her entire life no longer limited her. She raced her boys in the yard and jumped on the trampoline without fatigue or pain” ("An Answered Prayer: Warrior Reflects on Her Experience as the First Patient Treated with CRISPR for Sickle Cell Disease"). This innovative procedure was enabled by the remarkable properties of stem cells. Stem cells are the body's foundational cells, they are adaptive and capable of self-renewal being able to compose varying specialized cell types (e.g., cardiac muscle cells, red blood cells, neuron cells, thyroid cells, etc.), making them notably valuable in regenerative medicine. These complex eukaryotes, found in nearly every tissue in the human body, are exceptionally beneficial and versatile because they supply continuous tissue maintenance and regeneration. More specifically, stem cells heal the body through mitosis, with a portion of the resulting organisms remaining as stem cells and others differentiating into specific cell types needed to replace damaged or dead cells ("Stem Cells: What They Are and What They Do"). Stem cell differentiation is a process that is very highly regulated where pluripotent cells evolve into their specialized forms, driven through elaborate molecular processes, including “epigenetic modifications like DNA methylation and chromatin remodeling, which regulate gene expression and establish cell identity” (Cona). As Dr. Louis A. Cona more broadly relays, this process involves “the differential activation of specific genes and the repression of others, which changes the cell’s size, shape, function, and metabolic activity” (Cona). Genetic engineering-related treatment, like Victoria Gray’s, functions through enhancing stem cell remedies by modifying them to express therapeutic genes and correcting genetic defects at their source. Genes can be distinctly altered using precise tools like CRISPR-Cas9—a protein-specific DNA cleaver for CRISPR—which can cut DNA at pinpoint locations to add new genes, or remove or replace defective genes, thereby allowing the stem cells to emend (using differentiation) genetic defects or produce curative proteins (Rafiti et al.). Beyond CRISPR-Cas9, other gene-editing methods, such as ZFNs (Zinc Finger Nucleases) and TALENs (Transcription Activator-Like Effector Nucleases), also enable precise genetic modifications by adding, removing, or replacing genes to benefit the individual. Generally speaking, workings like these all must rely on engineered proteins to recognize and cut targeted DNA regions, facilitating the genetic alterations (Joung and Sander).

As genetic engineering advancements continue to transpire, the boundless possibilities of using it to treat and even eradicate numerous diseases and disorders are becoming increasingly feasible. For instance, gene editing technologies have the ability to, and could be used to alter traits such as height, physical appearance, or intelligence, in addition to curing diseases and defects. While physical attributes, notably height and appearance, already can be adjusted or manipulated based on current scientific knowledge and technology, intelligence is more far-fetched due to its polygenetic (relying on multiple gene contributions to synthesize the specified trait) nature (Carroll and Hyman). Despite this, it is increasingly rational to anticipate plausible intellectual enhancement in the near future due to ongoing scientific progress and emerging theories: “You could sequence the genomes of a couple’s embryos prior to implantation and select the one with the selection of SNPs most correlated to intelligence”

(Chivers). Significant ethical concerns regarding justice, consent, and the possibility of inequality must be addressed regarding the modification of human characteristics, though, most importantly intelligence. Allowing such changes could exacerbate already-existing social divisions by creating a society in which the wealthy have disproportionate access to genetic advantages. Reasonably, the ultimate goal of genetic engineering should continue to be the treatment of illnesses and genetic flaws, not the improvement of natural human characteristics, in order to avoid the moral conundrums and societal hazards that come with said developments. Genetic engineering should be lawfully limited to curing diseases and genetic defects, as using it for enhancements like intelligence could worsen social inequality and intensify moral issues about altering human nature. Additionally, all types of genetic engineering (apart from medical intervention) challenge ideas of fairness, as they could potentially create divisions in society based on access to enhancements, assuredly providing the possibility for undermining human values and undercutting the credit of human achievement, further justifying its practical need for prohibition.

Cognitive Genetic Engineering Ethics

Logically and ethically speaking, genetic engineering in the realm of enhancing human intelligence should be ruled out entirely because the social risks, distinguishably threatening autonomy and undermining the value of certain people, outweigh the potential benefits. Currently, there is still significant scientific uncertainty, notably with the polygenic nature of the brain, and lack of sufficient testing in the field: “We know the genes are in there, but because each one accounts for such a small proportion of the variance, they are hard to pinpoint” (Chabris et. al). However, even if, hypothetically, it was completely safe and reliable, having no major negative side effects associated with an enhancing procedure, the risks of developing a more discriminatory society and diminishing personal autonomy make the idea unjustifiable.

In a society without strict regulations on genetic engineering, wealthy individuals could, and would be likely to, exploit such technologies resulting in a distinct genetically advantaged class. Denying this premise as highly probable would be ignoring the consistent patterns of human behavior shown all throughout history. In fact, those in positions of power have frequently exploited their advantages to further their own interests, often at the expense of societal well-being and equity. For example, robber barons manipulated the economy during the 19th century in the United States for their own gain by exploiting labor, monopolizing industries, and using political influence to suppress competition. They consolidated wealth at the expense of others, employing ruthless tactics such as underpaying employees and using unfair pricing to maximize profits (Schneider). Similarly, during colonization in the 16th century, individuals in power exploited their position for personal gain by extracting resources from colonized regions, imposing economic systems that benefited the colonizers, while also creating institutions designed to retain control at the expense of the locals (Robinson and Daron). As shown in previous trends, it seems highly realistic for another incident, similar to colonization and robber baron exploitation, to occur under the proper circumstances. To sum up the particular case

regarding cognitive genetic alterations, "In an unregulated scenario, elites could use gene editing to enhance intelligence, physical traits, and longevity, creating a genetically superior class and exacerbating socio-economic and health disparities" (Youvan).

Next to this, cognitive genetic enhancement poses a serious threat to personal autonomy. Because modifying genes has a far more significant effect if applied in the early stages of human development, genetic modifications would often be made by parents before birth, where they have the ability to administer predetermined traits to their children which could limit their ability to make autonomous choices about their own identities and life paths. As philosopher Jonathon Pugh attests, genetic enhancement threatens autonomy by enabling external control over an individual's traits, limiting their capacity for self-determination and diminishing authentic personal identity (Pugh). To put this into perspective, in the hypothetical scenario people are genetically enhanced (mentally), it can be argued that people's traits, cognitive capacities, or abilities are theoretically predetermined through their modifications rather than conscious decisions or choices. In such a world, would people even be themselves, or would they function like machines, where they must conform to standards or boundaries set by their genetic ability rather than their own natural choices? True decision-making involves more than just intellectual power; it includes values, emotions, and individual experiences that shape preferences as well as actions. With enhanced cognitive abilities, people may be able to process information more quickly, however, genetic modifications could diminish freedom by adding a standardized cache of abilities and behaviors that influence, and possibly limit, the range of choices individuals feel they can make. To illustrate further, consider a situation where individuals feel as though they are programmed to function in a certain manner, much like a car constructed for a specific purpose like high-speed racing. Although the car is capable of other functions, society, as well as the car designers, expects it to race, anything else would seem like a waste of its potential. In the same fashion, a person who is cognitively enhanced would likely feel as though their only valid path is to pursue an intellectual field such as science, even if their true passion lies somewhere else completely, such as the arts or entrepreneurship. Conclusively, genetic enhancements in the realm of intelligence pose serious ethical and social risks, primarily exacerbating inequality and threatening personal autonomy. Because of these reasons, it is rational to rule out the concept of intellectual genetic engineering in order to benefit society and the individual well-being of humans.

Physical Alterations: Autonomy and Social Risks

Whilst cognitive genetic engineering is often viewed as more objectively problematic, it can be easier to argue that modifying physical traits might not present as significant a threat to individuals or society in comparison. Admittedly, enhancing physical traits seems harmless at first glance, not directly threatening major aspects of society or individuals' mental well-being. However, on further speculation, the concept of such alterations, like that of height or eye color, can have dire effects on society including reinforcing harmful standards and creating a superficial divide. As a general concept concerning genetic engineering for all organisms

(aligning with this specific notion), seemingly harmless modifications could have “unintended impacts on non-targeted organisms” (Kahn and Chan). Physical alterations to the human body via genetic engineering could largely be harmful to society by backing already baneful human beauty standards and creating marginalized groups as a result. With these specified beauty standards (e.g., tall height, symmetrical face, etc.), forms of discrimination or bias against individuals not possessing sought-after traits could become increasingly evident. Here, those not conforming to the standard would become marginalized (So). This, on a smaller scale, is already overtly obvious in society today with the observed postulation referred to as “pretty privilege”, where it is clinically proven that a positive light is shunned on more conventionally attractive people to the public: “wherein someone who is perceived positively in one aspect is assumed to possess other positive qualities as well” (Travers). Stemming from this phenomenon, individuals who fit these idealized standards are typically treated better and given more opportunities in various contexts, where they are not only more likely to be hired but also tend to earn higher salaries and receive more promotions (Travers). Agreeably, physical gene engineering would contribute to the issue of unfair advantages given to attractive people compared to others, and in turn corrupt society further in this aspect. Some may contend that since practices like plastic surgery and steroid usage already are practiced, and they too alter a person’s appearance or abilities, genetic engineering would not differ significantly from these existing methods. Nevertheless, while these procedures are undeniably problematic in their own sense, they remain temporary and oftentimes come with inherent health issues, societal stigma, as well as the need for continuative maintenance (Bonel et al.). Genetic engineering, on the other hand, fundamentally changes a person’s genetic makeup, creating permanent alterations to their body which is unable to be reversed. This permanence gives genetic engineering the opportunity to be far more damaging, as it introduces irreversible changes, unlike temporary procedures which allow for the possibility of reversal or correction. Reflecting on the key concerns, making one's physical appearance a major factor in determining one's success and treatment, if reinforced by genetic engineering, could exacerbate societal inequality significantly. Determinatively, society runs the risk of creating even more splits and further marginalizing those who do not fit these elevated standards, by manipulating characteristics to fit subjective beauty ideals. This ultimately would disrupt society and greatly undermine the value of certain people, making it illogical to permit genetic engineering in the domain of non-cognitive physical improvement, in addition to cognitive improvement.

Yet another concern when minding the genetic alteration of physical traits is the diminishment of human achievement. Humans, hypothetically, achieving feats (e.g., setting a long jump world record, deadlift world record, world record mile time, etc.), based on their artificial, genetically engineered ability rather than from their own effort and determination undermines foundational principles that drive societal growth. A shift so abrupt could very possibly change the way we as humans view achievement as a whole, with a transition from crediting human hard work, to crediting genetically engineered traits (Kahn and Chan). Theoretically, as a direct result of this phenomenon, individuals would not feel the motivation to

be persistent or work as hard, which undoubtedly harms society and the community. The underlying psychological motivation to reach one's goals can be driven by a sense of self-efficacy: the belief that one can influence outcomes through their own actions (Bandura). This feeling can be closely related to a feeling of pride or satisfaction resulting from overcoming challenges or personal achievement in general (Bandura). If genetic enhancements override this process, there's a risk that the intrinsic rewards of hard work would seriously decline, leading to reduced motivation to improve oneself and even an overall reduction in personal fulfillment (Bandura). Moreover, this effect could also cause a more passive society to develop, where fewer people take initiative or put in the work required to make meaningful contributions, feeling their achievements aren't attributable to their own efforts, but rather their artificial traits. Genetic engineering enhancing physical traits could undermine the value of hard work and achievement which would reduce motivation and self-efficacy. This idea could create a more passive society in which personal growth and societal progress are hindered, reinforcing the claim that genetic engineering should only be allowed for medical purposes, not physical betterment purposes as it weakens human achievement.

Limits of Personal Choice

Some argue that genetic enhancements should be a personal choice, as long as they don't directly affect others, but this view overlooks the broader implications of the matter. This libertarianism-centered perspective about society, the idea that "if you don't support a practice then simply do not participate", is overused and unrealistic. The fact of the matter is we live in a society where individual choices inevitably affect the collective. For this particular circumstance, even if one person's genetic modifications don't directly harm others, they still create an unfair advantage in the category of that trait (e.g., if someone is genetically enhanced to be stronger they inherently have a leg up in a weight lifting contest), as well as reap other societal repercussions because of their personal choice. Genetic engineering as a whole, outside of medicine, of course, creates a disparity between individuals, undermining fairness and equal opportunity. Genetic enhancements would fundamentally disrupt the balance of fairness, as not everyone would have access to the same modifications, leading to divisions based on who has the means to enhance themselves. All in all, it's not just about personal choice but rather the collective well-being of society.

Practical Restrictions

Altogether, while genetic engineering holds incredible potential for treating diseases, curing mutations, and alleviating human suffering, its use in enhancing human characteristics, specifically intelligence and other physical abilities, raises significant ethical and societal concerns, enough to the point it should be heavily restricted or disallowed by law. Allowing genetic modifications for personal improvements, such as increased intelligence or physical enhancements can worsen social inequality, form unfair advantages, and impair the value of hard work and human achievement. These modifications would not just sectionally divide society on

who can afford or access them but additionally threaten individual autonomy by imposing predetermined traits limiting identity and personal choice. The reality is, that we cannot ignore the broader ramifications of genetic engineering since every personal choice has a ripple effect on the collective well-being of society. To avoid creating a genetically divided living dynamic, genetic engineering should be strictly regulated or restricted entirely by law, limited to medical applications, and not used for enhancements that could harm the integrity of our society, fairness, and the principles of human achievement.

Works Cited

AABB. "An Answered Prayer: Warrior Reflects on Her Experience as the First Patient Treated with CRISPR for Sickle Cell Disease." AABB, 12 Sept. 2024, www.aabb.org/news-resources/news/article/2024/09/12/an-answered-prayer--warrior-reflects-on-her-experience-as-the-first-patient-treated-with-crispr-for-sickle-cell-disease.

Bonell, Sarah, et al. "The Cosmetic Surgery Paradox: Toward a Contemporary Understanding of Cosmetic Surgery Popularisation and Attitudes." PubMed, 2021, https://pubmed.ncbi.nlm.nih.gov/33962222/.

Chabris, Christopher F., et al. "Genetic Studies of Academic Achievement." Behavior Genetics, https://doi.org/10.1007/s10519-013-9609-0.

Cona, Louis A. "Stem Cell Differentiation and Regenerative Medicine." ScienceDirect, 2022, www.sciencedirect.com/ science/article/abs/pii/S0966327422001952.

Genoway. "CRISPR-Cas9 Gene Editing." Genoway, www.genoway.com/technologies/crispr-cas9-gene-editing.

Hunt, Dame Vivian, et al. "Why Diversity Matters." McKinsey & Company, 1 Jan. 2015, www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/why-diversity-matters.

Joung, J. Keith, and David R. Liu. "Genome Editing with CRISPR-Cas9." National Center for Biotechnology Information, 2013, www.pmc.ncbi.nlm.nih.gov/articles/PMC3547402.

Khan, Muhammad Aamir, and Ching Chan. "Ethical Considerations in Genetic Engineering." Plantae, 18 Mar. 2024, https://plantae.org/ethical-considerations-in-genetic-engineering/

Lopez-Garrido, Gabriel. "Bandura's Self-Efficacy Theory of Motivation in Psychology." Simply Psychology, 19 Feb. 2025, www.simplypsychology.org/self-efficacy.html

Mayo Clinic Staff. "Stem Cells: What They Are and What They Do." Mayo Clinic, 17 Dec. 2022, www.mayoclinic.org/tests-procedures/bone-marrow-transplant/in-depth/stem-cells/art-20048117.

McCarroll, Steven A., and Steven E. Hyman. "Progress in the Genetics of Polygenic Brain Disorders: Realizing Psychiatrically Relevant Insights from Large-Scale Genomic Studies." Neuropsychopharmacology, https://pmc.ncbi.nlm.nih.gov/articles/PMC4066986/.

Mind Matters. "Could We Really Increase Human IQ via Genetic Engineering?" Mind Matters, 22 Aug. 2022, www.mindmatters.ai/2022/08/could-we-really-increase-human-iq-via-genetic-engineering.

Pugh, Johnathon. "Autonomy, Rationality, and the Ethics of Genetic Enhancement." The American Journal of Bioethics,

https://pmc.ncbi.nlm.nih.gov/articles/PMC4347020/.
Robinson, James A., and Daron Acemoglu. "The Economic Impact of Colonialism." VoxEU, Centre for Economic Policy Research, 30 Jan. 2017, https://cepr.org/voxeu/columns/economic-impact-colonialism.
Schneider, Steven. "Robber Baron." Encyclopaedia Britannica, Encyclopaedia Britannica, Inc. 26 Mar. 2025, https://www.britannica.com/money/robber-baron.
So, Derek. "Human Germline Genetic Modification and Human Dignity: An Ethical Analysis." PubMed Central, National Institutes of Health, 7 July 2021, https://pubmed.ncbi.nlm.nih.gov/34218455/
Travers, Mark. "A Psychologist Explains the 'Pretty Privilege' Paradox." Forbes, 15 Oct. 2023, www.forbes.com/sites/traversmark/2023/10/15/a-psychologist-explains-the-pretty-privilege-paradox/.
Youvan, Douglass C. The Ethics of Genetic Enhancement: Examining the Societal Consequences of Unregulated Gene Editing for the Elite. ResearchGate, October 2024, https://www.researchgate.net/publication/384598399_The_Ethics_of_Genetic_Enhancement_Examining_the_Societal_Consequences_of_Unregulated_Gene_Editing_for_the_Elite

Empowering Families Through Genetic Testing: The CFTR Gene and Cystic Fibrosis
By Amelia Weldon

Advances in genetic testing have revolutionized modern healthcare, especially for families where one parent is known to carry a mutation. In the case of cystic fibrosis (CF), caused by mutations in the *CFTR* gene, carrier screening can illuminate the risk of transmission, guiding vital decisions and support systems. The implications of being a carrier extend far beyond personal health—they shape family planning, emotional preparedness, and medical strategies for future generations. Genetic testing—especially CFTR carrier screening—is not only the basis of informed family planning, but also improves early intervention, emotional readiness, and broader familial awareness, fundamentally enhancing outcomes for families facing cystic fibrosis.

One of the most compelling benefits of CFTR genetic testing is its impact on family planning. A carrier typically has one normal gene and one mutated gene for a recessive condition, meaning they do not exhibit symptoms but can pass the mutation to their offspring. If the other parent is tested and also found to be a carrier, the risk of having a child with CF is 25% with each pregnancy ("Carrier Screening for Genetic Conditions"). This knowledge enables prospective parents to make informed decisions about conception, whether through natural means, assisted reproductive technologies like in vitro fertilization with preimplantation genetic testing, or alternatives such as adoption. Access to accurate risk data transforms reproductive decisions from guesswork into proactive planning (MyNucleus).

Beyond reproductive choices, carrier screening provides families with a clear and quantifiable assessment of risk. Carrier testing—whether through saliva or blood—identifies specific CFTR mutations and estimates the likelihood of transmission ("What Is Carrier Screening?"). The American College of Obstetricians and Gynecologists recommends offering CF carrier screening to all individuals who are pregnant or planning a pregnancy to ensure they have the information needed to make fully informed decisions ("Carrier Screening for Genetic Conditions"). This clarity fosters candid, evidence-based discussions between parents, healthcare providers, and genetic counselors.

If a child is at risk or diagnosed with cystic fibrosis, early intervention can significantly improve outcomes. Since the mid-2000s, all U.S. states have implemented newborn screening programs for CF ("Newborn Screening for CF"). These tests detect elevated immunoreactive trypsinogen levels and can identify mutations in the CFTR gene before symptoms arise. Early diagnosis allows for immediate initiation of treatments such as airway clearance techniques, nutritional support, and—more recently—CFTR modulator therapies. Studies indicate that early CFTR modulator treatment can slow disease progression and improve lung function and nutritional status (Coverstone et al.). In some cases, intervention within the first months of life can alter the disease trajectory entirely.

While genetic testing offers clinical advantages, it also plays a vital role in emotional preparedness. Parents who know they carry genetic risks can mentally and logistically prepare

for potential challenges. This may include arranging for specialized medical care, seeking out community support, or accessing counseling services to manage stress and anxiety. Genetic counseling provides a bridge between complex genetic data and personal decision-making, guiding families through the emotional, ethical, and medical dimensions of their results ("Genetic Counseling"). Preparedness does not eliminate the difficulties of raising a child with a chronic illness, but it can reduce uncertainty and help families feel more in control of their circumstances.

The benefits of CFTR genetic testing also ripple outward into the extended family. When one individual is identified as a carrier, it is possible that siblings, cousins, and other relatives may also carry the same mutation. This awareness can prompt relatives to seek testing, creating opportunities for informed reproductive decisions across generations ("Information for Carriers"). Such "cascade screening" fosters a family culture of shared health knowledge and can lead to earlier detection of conditions in other family members.

Finally, genetic testing supports a move toward personalized medicine. For children at risk of CF, healthcare providers can design individualized monitoring and treatment plans. This might involve frequent lung function assessments, targeted nutritional interventions, and preventative care to reduce infections (Centers for Disease Control and Prevention). By tailoring care to the patient's specific genetic and clinical profile, healthcare providers can improve both quality of life and long-term prognosis.

In conclusion, CFTR genetic testing for families where one parent is a carrier is far more than a medical formality—it is an essential component of modern preventive healthcare. From informed reproductive decisions and precise risk assessment to early interventions, emotional readiness, and generational awareness, the benefits of genetic testing are clear and far-reaching. As science advances and therapeutic options expand, the value of identifying genetic risk will only grow. Embracing genetic testing is not simply about knowing one's DNA—it is about empowering families to make decisions that lead to better health outcomes and more resilient futures.

Works Cited

"Carrier Screening for Genetic Conditions." Obstetrics & Gynecology, Committee Opinion No. 691, American College of Obstetricians and Gynecologists, Mar. 2017. ACOG, www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2017/03/carrier-screening-for-genetic-conditions.

Centers for Disease Control and Prevention. "Newborn Screening for Cystic Fibrosis: Evaluation of Benefits and Risks and Recommendations for State Newborn Screening Programs." MMWR Recommendations and Reports, vol. 53, no. RR-13, 2004, www.cdc.gov/mmwr/preview/mmwrhtml/rr5313a1.htm.

Coverstone, Allison M., et al. "Early Diagnosis and Interventions Have Led to Improved Clinical Outcomes for People with Cystic Fibrosis." Pediatric Clinics of North America, vol. 68, no. 6, 2021, pp. 1261-1280. PMC, pmc.ncbi.nlm.nih.gov/articles/PMC7830672/.

"Genetic Counseling." Wikipedia, Wikimedia Foundation, en.wikipedia.org/wiki/Genetic_counseling.

"Information for Carriers." Cystic Fibrosis Trust, www.cysticfibrosis.org.uk/what-is-cystic-fibrosis/diagnosis/information-for-carriers.

MyNucleus. "Cystic Fibrosis Carrier Testing." MyNucleus, 29 Apr. 2025, mynucleus.com/blog/cystic-fibrosis-carrier-testing.

"Newborn Screening for CF." Cystic Fibrosis Foundation, www.cff.org/intro-cf/newborn-screening-cf.

"What Is Carrier Screening?" Verywell Family, www.verywellfamily.com/what-is-carrier-screening-5235603.

Alzheimer's Disease: Biological Origins, Diagnostic Tools, Challenges and Recent AI Developments By Mia Luh

Abstract

Alzheimer's disease (AD) is a progressive brain disorder that impairs memory and cognitive function, largely due to the buildup of amyloid plaques and neurofibrillary tangles. These abnormal protein accumulations interfere with communication between neurons, leading to their gradual deterioration—a process that researchers are now exploring new ways to diagnose and treat, including the use of artificial intelligence (AI). Previous studies have used brain imaging techniques, such as Positron Emission Tomography (PET) scans and Magnetic Resonance Imaging (MRIs), to track the accumulation of amyloid plaques and neurofibrillary tangles in Alzheimer's patients, aiming to understand how these abnormalities contribute to cognitive decline. This paper examines their roles in AD pathogenesis and explores how artificial intelligence (AI) can enhance early detection and treatment. More recently, AI-driven models have been developed to analyze brain scans and patient histories, improving early diagnosis and predicting disease progression. Building on these methods, my research will first characterize the roles of amyloid plaques and neurofibrillary tangles in AD pathogenesis, then explore how AI can leverage existing neurological data to enhance diagnostic accuracy and identify novel treatment strategies—ultimately enabling earlier interventions to slow cognitive decline. Ultimately, this approach highlights how AI may transform Alzheimer's care and improve patient outcomes. (201)

Keywords

Alzheimer's Disease, Neurofibrillary Tangles, Amyloid Plaques, Tau Protein, Alpha Secretase, Beta Secretase, Gamma Secretase, Artificial Intelligence

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia, accounting for 60–70% of all dementia cases worldwide (WHO). Over 55 million people currently live with dementia, a number projected to triple to 139 million by 2050 due to aging populations (Alzheimer's Disease International). In the U.S. alone, 1 in 9 adults over 65 has AD, with total annual costs exceeding $345 billion (Alzheimer's Association). Much research is being conducted into the cause and mechanism of action of the disease, however, confirmed information is still not readily available (NIH). Although the theory of Alzheimer's has been prevalent since the Ancient Egyptians, the disease was only coined until 1910 in a publication by Alois Alzheimer. He studied a middle-aged woman—Auguste Deter (1850-1906)—who displayed signs of uncontrollable behavior, loss of memory, and delusion. Through a biopsy of her brain, Alzheimer revealed significant atrophy within her cerebral cortex alongside finding amyloid plaques and neurofibrillary tangles (Yang et al.). As research

continued, more details about the pathology of Alzheimer's were discovered, particularly the formation of amyloid plaques and neurofibrillary tangles.

In this paper, I will first review the biological origins and key pathological features of Alzheimer's Disease. Then, I will examine current diagnostic methods, highlighting their strengths and their limitations. Finally, I will explore recent advances in Artificial Intelligence as they relate to Alzheimer's Disease, discussing how these innovations address existing challenges and proposing directions for future areas of exploration.

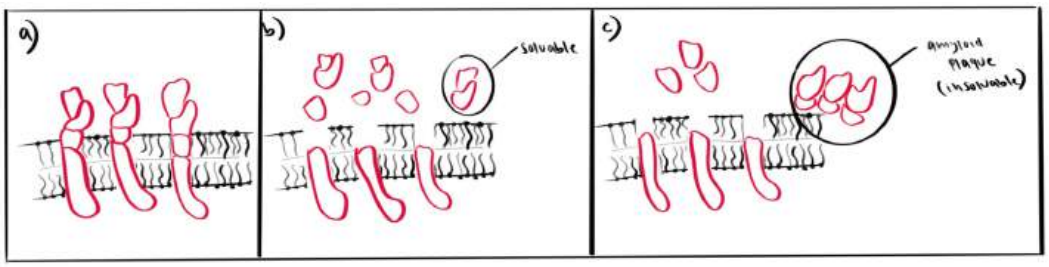


Figure 1: A visualisation of Amyloid Precursor Protein (APP) and its cleavage.
A. APP at rest in the cell membrane. B. **Non-Alzheimer's.** The protein is cleaved in two soluble parts.
C. **Alzheimer's.** *Gamma* secretase works with *beta* secretase to cleave beta-amyloid (non-soluble). (Mia Luh)

Molecular mechanisms

Alzheimer's disease arises from complex molecular disruptions that progressively impair neuronal function, primarily driven by amyloid-β (Aβ) plaque accumulation and tau-mediated neurofibrillary tangles (NFTs). These pathological processes, exacerbated by genetic and environmental risk factors, trigger synaptic dysfunction, neuroinflammation, and irreversible neurodegeneration. The plaques are formed by recycled parts of the Amyloid Precursor Protein (APP) extending both inside and outside of the cell (Fig. 1a). In a healthy cell membrane, once APP is utilized and broken down, it gets cleaved by alpha and gamma secretase and is divided up such that the removed parts are soluble in the extracellular fluid (Fig. 1b). However, if the gamma secretase instead works with a beta secretase, it creates an insoluble fragment called amyloid-beta (Gouras et al.). These fragments bind together to create amyloid plaques (Fig. 1c); these plaques disrupt neuron communication by lodging between neurons rendering them unable to signal. Amyloid plaques can also congregate by the blood vessels in the brain increasing the chance of hemorrhage and blood loss.

The amyloid plaque also triggers a separate reaction that triggers phosphates to bind to tau proteins. A neurofibrillary tangle is the accumulation of the protein tau inside the neuron. Without tau to support the microtubules, the neuron is not able to signal properly, therefore contributing to loss of neuron communication (Medeiros et al.). In some cases, the neuron will undergo apoptosis causing high levels of cell death in the brain. Apoptosis can be seen as cell shrinkage, increased extracellular water, or altered diffusion properties (see Fig. 2 for example

images of this effect at different severity levels). Intense cell death in the brain can affect the localized function of each area leading to symptoms such as memory loss or confusion.

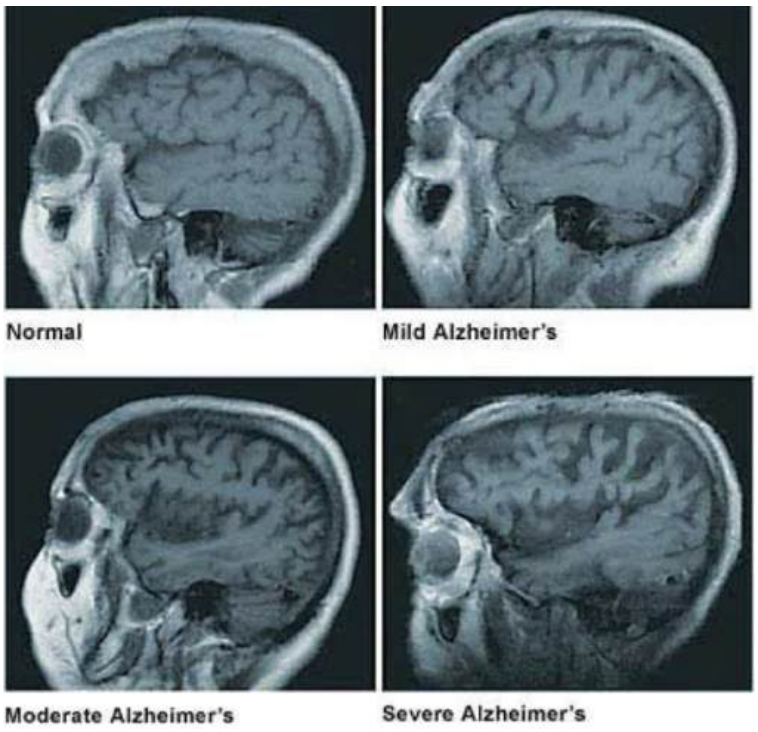


Figure 2. **Progression of brain atrophy in Alzheimer's disease.**
Sagittal MRI scans showing progressive brain shrinkage from a healthy brain (A) to mild (B), moderate (C), and severe (D) stages of Alzheimer's disease. As the disease advances, there is visible cortical thinning and enlargement of the ventricles due to neuronal loss. Hippocampal atrophy becomes more pronounced in later stages, correlating with memory impairment and cognitive decline.
Adapted from Mayo Clinic (n.d.),
https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/multimedia/brain-atrophy-and-alzheimers-/img-20008229

Genetic risk factors

Understanding the mechanism of action of Alzheimer's has also exposed the potential genetic risk factors. Most Alzheimer's cases are idiopathic, with a higher likelihood occurring in individuals over 65. The exact cause of Alzheimer's is unknown, however it is speculated that lifestyle, age, and environment play a role in its development, although small genes have been found to impact the likelihood of developing the disease as well. There are currently 80 genetic regions associated with Alzheimer's yet only 3 have been identified as sources of the disease (Alzheimer's Disease Fact Sheet, 2023), for instance, the e4 apolipoprotein (APoE e4) gene. There are two versions of the gene which can be inherited, e4 or e2. Inheriting one e4 allele increases the chances of developing Alzheimer's while inheriting two, further escalates the risk. The apolipoprotein is responsible for breaking down beta-amyloid and the e4 allele is less effective than the e2 allele.

Other examples of genetic influence are the PSEN-1 or PSEN-2 mutations (Alzheimer's Disease Fact Sheet). These are responsible for where the gamma secretase cleaves APP which could lead to larger fragments gathering more easily. Lastly, individuals with down syndrome

have an extra chromosome 21, which encodes for APP. This results in an increased expression of APP, which can ultimately lead to an abundance of APP buildup (Alzheimer's Disease Fact Sheet).

Current Technology in AD Detection

Detecting Alzheimer's disease (AD) presents numerous challenges that complicate early and accurate diagnosis. One major obstacle is the subtlety of early symptoms, which often mimic normal aging or other neurological conditions, leading to misdiagnosis or delayed detection (NIH).

Currently, Alzheimer's diagnosis primarily relies on neuroimaging techniques such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and Computed Tomography (CT). These tools allow clinicians to observe structural changes in the brain, particularly the shrinkage of the hippocampus and buildup of amyloid plaques. Among these, Positron Emission Tomography (PET) scans are considered one of the most definitive tools because they can directly visualize the protein accumulations of Alzheimer's: amyloid-beta plaques and tau neurofibrillary tangles (Villemagne et al.). MRI scans can show hippocampal atrophy (see Fig. 2 for an example), a key area for memory and one of the first areas to be impacted by AD. A CT can show physical changes in the brain that are linked to AD, however it is not a definitive diagnostic tool. Although CT scans can visualize brain shrinkage or enlarged vesicles, they are primarily used to rule out other possible conditions. Additionally, cerebrospinal fluid (CSF) analysis helps detect levels of amyloid-beta and tau proteins, offering biochemical evidence of Alzheimer's pathology (Li et. al.).

However, the above brain imaging methods can be invasive, expensive, and inaccessible in many settings. On the other hand, cognitive assessments, while commonly used, introduce subjectivity as they depend on patient cooperation and clinician interpretation. The biological complexity of AD, with its varying progression patterns among individuals, makes it difficult to establish a universal diagnostic criteria. These challenges highlight the urgent need for more reliable, affordable, and accessible diagnostic tools to improve early detection and intervention for AD.

This is where Artificial Intelligence (AI) offers transformative potential. Rather than replacing current diagnostics, AI can enhance them by analyzing scan data more precisely and rapidly than human interpretation alone. For instance, machine learning algorithms can detect subtle variations in brain structure from MRI scans or identify characteristic patterns of amyloid deposition in PET scans that may be imperceptible to radiologists (Shimron and Perlman). Additionally, by integrating longitudinal patient data, including genetic profiles, cognitive tests, and imaging results, AI models can form a comprehensive and individualized assessment of AD risk (Fisher, et. al.). This bridging of biological understanding and computational modeling sets the stage for earlier and more accurate diagnosis, particularly before irreversible damage occurs.

Advancements in AI for Early Prediction

AI has revolutionized the detection of Alzheimer's disease (AD) by enhancing accuracy, enabling early diagnosis, and offering non-invasive solutions (Qiu et al.). AI techniques, particularly machine learning (ML) and deep learning (DL), analyze complex datasets such as brain scans, retinal images, and biomarkers to identify AD-related patterns with high precision (Ashayeri et al.). For instance, DL models like Convolutional Neural Networks (CNNs) can detect subtle structural changes in MRI scans that may indicate early-stage AD (Farooq et al.), while Long Short-Term Memory (LSTM) networks predict disease progression by analyzing longitudinal multimodal imaging data (Hong et al.). AI also addresses limitations of traditional methods by reducing reliance on subjective cognitive tests and improving sensitivity of imaging and CSF analyses in identifying biomarkers like amyloid-beta and tau proteins (Jack et al.). Additionally, AI algorithms applied to retinal imaging (Ashayeri et al.) and functional near-infrared spectroscopy (Hock et al.) recordings have been shown to identify AD with high accuracy, thus they can offer portable and cost-effective screening options, making early detection more accessible. Beyond diagnosis, AI aids in drug discovery (Qiu et al.) and personalized treatment planning (Fisher et al.), further advancing AD research and care. As AI technology continues to evolve, its integration into healthcare promises to transform AD detection and management on a global scale.

Discussion

Challenges and Ethical Considerations

While promising, AI-driven approaches face limitations due to the need for large, diverse datasets that are often hindered by privacy concerns and data fragmentation. The use of artificial intelligence (AI) in Alzheimer's disease diagnosis and treatment presents several ethical challenges that must be addressed to ensure responsible implementation.

One major concern is data privacy, as AI systems require large amounts of sensitive medical data such as brain scans and genetic information, raising questions about consent—particularly for patients with cognitive impairment who may not be able to provide informed consent themselves.

Additionally, AI models trained on non-diverse datasets can reflect and perpetuate existing healthcare disparities, leading to biased outcomes for underrepresented populations (Obermeyer et al.).

The lack of transparency in many AI systems, often referred to as the "black box" problem, makes it difficult for clinicians to understand and trust AI-generated diagnoses, potentially undermining their medical judgment (London).

Furthermore, accountability becomes complicated when AI tools contribute to clinical decisions—raising questions about who is liable if the AI errs: the developer, physician, or healthcare institution (Yu et al.).

Importantly, AI should be used to augment rather than replace human expertise to preserve patient autonomy and ensure ethical, patient-centered care. Ensuring equity, explainability, and oversight in AI applications is essential as these technologies become more integrated into Alzheimer's care.

Future Directions

As AI continues to evolve, the future of Alzheimer's research and care will likely benefit from greater personalization, accessibility, and prevention. One promising area for future research is the development of AI systems capable of combining various types of input—genetic data, neuroimaging, speech patterns, retinal scans, and even wearable biosensor data—into a unified diagnostic framework. This integrative approach could catch AD before clinical symptoms manifest, enabling pre-symptomatic intervention.

Furthermore, federated learning is gaining traction as a method to train AI models across decentralized databases without compromising patient privacy (Yurdem et al.). This technique allows AI systems to learn from global patient data while preserving confidentiality—helping solve the current limitations around data sharing.

Future AI models must be trained on diverse, representative datasets to reduce healthcare disparities. This involves efforts to collect data from underrepresented populations across different ethnicities, socioeconomic backgrounds, and geographic regions. Additionally, AI fairness auditing tools and bias-mitigation algorithms (Pfohl et. al,) can be integrated into the development process to continually assess and correct for bias before models are deployed in clinical settings.

Improving transparency requires the advancement of explainable AI methods (Ghassemi M et. al) that make machine learning outputs interpretable to clinicians. Tools like heatmaps for imaging AI and feature importance scores for diagnostic algorithms can help doctors understand why a particular prediction was made.

Clear legal frameworks and clinical guidelines are needed to define responsibility when AI is involved in decision-making (Naik et. al). Future steps include establishing AI validation standards similar to those used for medical devices, requiring rigorous clinical trials before deployment. Institutions should also implement clear accountability protocols, specifying when AI is used as a decision-support tool versus when it drives clinical action. Ultimately, responsibility will likely remain shared between healthcare providers and developers

Rather than replacing clinicians, AI should be developed as a clinical decision support system (CDSS) that enhances human expertise. Future training programs will likely incorporate AI literacy for healthcare professionals, empowering them to use AI responsibly while maintaining control over patient care.

Conclusion

Alzheimer's disease remains one of the most complex neurodegenerative disorders, caused by the accumulation of amyloid plaques and neurofibrillary tangles that disrupt neuronal

communication and function. While biological understanding of these marks has advanced significantly, challenges in early and accurate diagnosis persist—delaying intervention and treatment. Artificial intelligence recently emerged as a transformative solution, offering improved detection through the analysis of medical imaging, genetic data, and clinical histories. AI not only enhances existing diagnostic technologies but also opens doors to earlier identification, personalized treatments, and accelerated drug development. However, its integration into clinical care must be approached ethically, ensuring transparency, equity, and patient autonomy. By combining deep biological insight with cutting-edge computational tools, the future of Alzheimer's care holds the potential for more precise, accessible, and humane outcomes.

Works Cited

Alzheimer's Disease Drug Development Pipeline: 2024 - Cummings - 2024 - Alzheimer's & Dementia: Translational Research & Clinical Interventions - Wiley Online Library, alz-journals.onlinelibrary.wiley.com/doi/10.1002/trc2.12465. Accessed 3 Aug. 2025.

"Alzheimer's Disease Facts and Figures." Alzheimer's Association, www.alz.org/alzheimers-dementia/facts-figures. Accessed 3 Aug. 2025.

Alzheimer's Disease Fact Sheet | National Institute on Aging, www.nia.nih.gov/health/alzheimers-and-dementia/alzheimers-disease-fact-sheet. Accessed 3 Aug. 2025.

Hamidreza Ashayeri, et al. "Retinal Imaging and Alzheimer's Disease: A Future Powered by Artificial Intelligence." Graefes Archive for Clinical and Experimental Ophthalmology, 15 Feb. 2024, https://doi.org/10.1007/s00417-024-06394-0.

Betul Yurdem, et al. "Federated Learning: Overview, Strategies, Applications, Tools and Future Directions." Heliyon, vol. 10, no. 19, 1 Sept. 2024, pp. e38137–e38137, www.sciencedirect.com/science/article/pii/S2405844024141680, https://doi.org/10.1016/j.heliyon.2024.e38137.

"Brain Atrophy and Alzheimer's." Mayo Clinic, Mayo Foundation for Medical Education and Research, www.mayoclinic.org/diseases-conditions/alzheimers-disease/multimedia/brain-atrophy-and-alzheimers-/img-20008229. Accessed 3 Aug. 2025.

DeTure, Michael A., and Dennis W. Dickson. "The Neuropathological Diagnosis of Alzheimer's Disease - Molecular Neurodegeneration." BioMed Central, BioMed Central, 2 Aug. 2019, molecularneurodegeneration.biomedcentral.com/articles/10.1186/s13024-019-0333-5.

Education, MCL. "Alzheimer's Disease CSF Biomarkers." Insights, 6 Feb. 2025, news.mayocliniclabs.com/2022/02/07/alzheimers-disease-csf-biomarkers/.

Esteva, Andre, et al. "A Guide to Deep Learning in Healthcare." Nature News, Nature Publishing Group, 7 Jan. 2019, www.nature.com/articles/s41591-018-0316-z.

A Deep CNN Based Multi-Class Classification of Alzheimer's Disease Using MRI | IEEE Conference Publication | IEEE Xplore, ieeexplore.ieee.org/document/8261460/. Accessed 3 Aug. 2025.

Fisher Daniel W., et al. "Sexual Dimorphism in Predisposition to Alzheimer's Disease." Neurobiology of Aging, U.S. National Library of Medicine, pubmed.ncbi.nlm.nih.gov/29754747/. Accessed 3 Aug. 2025.

Fisher, Charles K., et al. "Machine Learning for Comprehensive Forecasting of Alzheimer's Disease Progression." Nature News, Nature Publishing Group, 20 Sept. 2019, www.nature.com/articles/s41598-019-49656-2.

Ghassemi, Marzyeh, et al. "The False Hope of Current Approaches to Explainable Artificial Intelligence in Health Care." The Lancet Digital Health, vol. 3, no. 11, 1 Nov. 2021, pp.

e745–e750, www.sciencedirect.com/science/article/pii/S2589750021002089, https://doi.org/10.1016/S2589-7500(21)00208-9.

Gouras, Gunnar K., et al. “β-Amyloid Peptides and Amyloid Plaques in Alzheimer’s Disease.” Neurotherapeutics, vol. 12, no. 1, 5 Nov. 2014, pp. 3–11, www.ncbi.nlm.nih.gov/pmc/articles/PMC4322079/, https://doi.org/10.1007/s13311-014-0313-y.

Hock, C, et al. “Near Infrared Spectroscopy in the Diagnosis of Alzheimer’s Disease.” Annals of the New York Academy of Sciences, vol. 777, 1996, pp. 22–9, pubmed.ncbi.nlm.nih.gov/8624087/, https://doi.org/10.1111/j.1749-6632.1996.tb34397.x.

Hong, Shenda, et al. “Opportunities and Challenges of Deep Learning Methods for Electrocardiogram Data: A Systematic Review.” Computers in Biology and Medicine, vol. 122, July 2020, p. 103801, https://doi.org/10.1016/j.compbiomed.2020.103801.

Jack, Clifford R, et al. “Tracking Pathophysiological Processes in Alzheimer’s Disease: An Updated Hypothetical Model of Dynamic Biomarkers.” The Lancet Neurology, vol. 12, no. 2, Feb. 2013, pp. 207–216.

Kumar, Anil, et al. “Alzheimer Disease.” National Library of Medicine, StatPearls Publishing, 12 Feb. 2024, www.ncbi.nlm.nih.gov/books/NBK499922/.

Li, Wentao, et al. “Alzheimer Disease Cerebrospinal Fluid Biomarkers in a Tertiary Neurology Practice.” Mayo Clinic Proceedings, 1 June 2024, https://doi.org/10.1016/j.mayocp.2023.12.024. Accessed 4 Sept. 2024.

London, Alex John. “Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability.” Hastings Center Report, vol. 49, no. 1, Jan. 2019, pp. 15–21, https://doi.org/10.1002/hast.973.

Medeiros, Rodrigo, et al. “The Role of Tau in Alzheimer’s Disease and Related Disorders.” CNS Neuroscience & Therapeutics, vol. 17, no. 5, 14 June 2010, pp. 514–524, https://doi.org/10.1111/j.1755-5949.2010.00177.x.

Naik, Nithesh, et al. “Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?” Frontiers in Surgery, vol. 9, no. 862322, 13 Mar. 2022, pp. 1–6. Frontiers, www.frontiersin.org/articles/10.3389/fsurg.2022.862322/full, https://doi.org/10.3389/fsurg.2022.862322.

National Institute on Aging. “How Is Alzheimer’s Disease Diagnosed?” National Institute on Aging, 8 Dec. 2022, www.nia.nih.gov/health/alzheimers-symptoms-and-diagnosis/how-alzheimers-disease-diagnosed.

Obermeyer, Ziad, et al. “Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations.” Science, vol. 366, no. 6464, 25 Oct. 2019, pp. 447–453, www.science.org/doi/10.1126/science.aax2342, https://doi.org/10.1126/science.aax2342.

Pfohl, Stephen R., et al. “An Empirical Characterization of Fair Machine Learning for Clinical Risk Prediction.” Journal of Biomedical Informatics, vol. 113, Jan. 2021, p. 103621, https://doi.org/10.1016/j.jbi.2020.103621. Accessed 5 Feb. 2021.

Qiu, Yunguang, and Feixiong Cheng. “Artificial Intelligence for Drug Discovery and Development in Alzheimer’s Disease.” Current Opinion in Structural Biology, vol. 85, 1 Apr. 2024, p. 102776, www.sciencedirect.com/science/article/pii/S0959440X24000034, https://doi.org/10.1016/j.sbi.2024.102776.

Shimron, Efrat, and Or Perlman. “AI in MRI: Computational Frameworks for a Faster, Optimized, and Automated Imaging Workflow.” Bioengineering, vol. 10, no. 4, 1 Apr. 2023, p. 492, www.mdpi.com/2306-5354/10/4/492, https://doi.org/10.3390/bioengineering10040492.

World Health Organization. “Dementia.” World Health Organization, 31 Mar. 2025, www.who.int/news-room/fact-sheets/detail/dementia.

Yang, Hyun Duk, et al. “History of Alzheimer’s Disease.” Dementia and Neurocognitive Disorders, vol. 15, no. 4, 31 Dec. 2016, pp. 115–121, www.ncbi.nlm.nih.gov/pmc/articles/PMC6428020/, https://doi.org/10.12779/dnd.2016.15.4.115.

Yu, Kun-Hsing, et al. “Artificial Intelligence in Healthcare.” Nature Biomedical Engineering, vol. 2, no. 10, Oct. 2018, pp. 719–731, https://doi.org/10.1038/s41551-018-0305-z.

The Current State of AI Music Generation By Ethan Kou

Abstract

This paper seeks to provide a comprehensive comparison between human-written and Artificial Intelligence (AI)-generated music. First, it explains the fundamentals of generative AI, before focusing in on the techniques currently used to generate music. After clearly defining what the AI can and cannot do, it then explains current casual and critical perceptions of the technology, including a survey to track how easily people can detect AI music. Finally, it delves into the ethical implications of the technology to see what everything defined up until then means in the real world.

1. Introduction

As generative AI becomes more and more popular, artists across all fields have needed to inform themselves of the tool's capabilities and limitations. However, such information is not widely distributed, and the way AI generates media is highly variable and frequently either partially or completely misunderstood. By investigating prior AI research and expanding on how AI differs from human creation, this paper disambiguates the murky waters surrounding generative AI.

1.1 An Introduction of Generative AI

The fundamental challenge in the field of artificial intelligence lies in enabling machines to emulate human-like intelligent reasoning. Despite their computational speed and efficiency, conventional computers remain limited to explicitly programmed, highly specific tasks. This makes autonomous cognition like that of a human an inherently complex and yet unsolved problem.

Hence to make AI, an engineer must instead get a computer to act as close to intelligent as possible for an extremely specific task. This approach is the one used by generative AI systems, operating under constrained parameters to simulate human-like output. By analyzing a large set of data, AI can employ statistical modeling to synthesize content with characteristics that could be attributed to human creation, thereby achieving a convincing facsimile of creativity.

1.2 How LLMs work

At its most basic, this process of replicating human behavior is best demonstrated by one of the most popular AI models, the Large Language Model. The LLM is at the center of text generating and predicting AIs like ChatGPT, and provides an excellent introduction to several fundamental AI concepts.

The first such concept is training. In essence, the LLM is one large mathematical function that repeatedly predicts the word most likely to come next in any given sentence. To do this, it must first be trained on a massive amount of sentences, all of which contain varying words at various frequencies in a variety of patterns of structures. Training consists of feeding the AI

these sentences word by word, comparing the actual next word to its predicted output and changing the likelihood of that word being chosen the next time a similar sentence appears. Upon completion, the AI now has a long list of probabilities it can rely on to determine the next most likely word in almost any situation, allowing it to complete new sentences and generate its own.

Tokenization is also vital to LLMs and to AI music generation. In an LLM, tokens are stored as long vectors of numbers that encode the "meaning" of each word. It is helpful to imagine each vector like a strand of DNA, and each number like a protein bond that stores vital information on how the word it tokenizes is being used. By transforming a sentence into a sequence of tokens, the LLM is capable of attention, a process where each token is changed based on the other tokens around it. A good example would be a token encoding the word "bank". By itself, the word could either mean a riverbank or a savings bank, but through attention, the word's true meaning is encoded based on the context of the surrounding sentence. If there is a "money" token nearby, the "bank" token's number would change to reflect that specific meaning, and the same would go for if there were a "river" token.

By repeatedly running attention processes followed by comparing the sentence to its training through a feedforward process, the LLM is able to distill the most essential meaning of a sentence in order to decide its next words with the utmost accuracy. [1]

1.3 How AI Generates and Recreates Sound

An in-depth understanding of LLMs and tokenization is necessary for a seamless explanation of Google's AudioLM, an AI sound generation model. Working in much the same way as an LLM, we can use AudioLM as our model for high quality sound generation.

AudioLM's tokenization is split into a hierarchy of three levels, to account for the much higher data rate of audio compared to words. The first level is semantic tokens, which capture the musical patterns in broad strokes, encoding things like tone, instrumentation, and vocals into its vectors. The second and third levels are composed of coarse and fine acoustic tokens, which create consistent detailing in the sound. They ensure the audio has a consistent quality and fill in any of the gaps left behind during semantic tokenization. By creating these tokens separately and then stitching them back together into a complete recording, AudioLM is able to predict and create audio clips much like LLM does for sentences. [2]

1.4 Introducing Suno AI Model

This paper focuses on the AI music generator Suno, one of the many products that have come out using AudioLM as a template for generating a wide variety of music. It uses an AI model called Bark to create lyrics and songs using the same 3-tiered tokenization hierarchy as AudioLM. [3] On the Suno website, users can describe the piece they want to create and provide tags that directly influence what semantic tokens the model generates.

2. What's Missing in AI Music

After thorough analysis and experimentation with the model, I identified key patterns and limitations in its song generation. Some of these limitations have to do with the way Suno is trained. It allegedly uses “all music files of reasonable quality that are accessible on the open internet” [4]. This is not a completely random sample of all music everywhere, and thus severely limits the kinds of music Suno can competently reproduce. For instance, Suno’s instrumentation defaults to guitars and pianos, the most common instruments found across all free-access music, limiting its ability to generate sounds from other instruments. Additionally, a substantial portion of the music Suno generates is reminiscent of pop due to the genre composing the vast majority of songs it has access to.

However, several flaws can be attributed to the AI itself, most glaringly the motifs it creates and how they connect to tokenization. A musical motif, a recurring pattern or melody that identifies the song and its core emotions, is at the center of almost every song in existence. At first glance, the repetitive variation of motifs would seem like a place where the AI would excel in creating music. However, taking that approach and seeing how it plays out over the course of a full-length song reveals the weaknesses of musical tokenization in songwriting. Across all songs generated by Suno AI, motifs are distinctly shorter, more repetitive, and more predictable than those of human artists. Each song loops the motif repeatedly with minimal variation and provides very little to suggest meaningful progression or escalation.

While on a micro level, the AI can convincingly recreate real sounds and make them close to legitimate music, on the macro scale, AI music has yet to make a complete composition that brings out anything deeper. It is limited by its tokens, which, by nature, have to be small and short to contain data, but their brevity fails to create a bigger picture of the song they encode.

3. Perception of AI Music

It is important to see how music generated by AI is actually perceived by the general public. To do this, we created a survey that asked people to guess which of two songs were generated by AI. The survey asked this question about four pairs of songs, and then asked respondents to rate them in terms of emotional, aesthetic, and creative quality.

Since AI generated songs were shorter on average, I cut most of the human generated songs down to a matching length, ending each one at a sensible point. Most of these snippets ended up around two minutes, though some longer AI songs pushed their lengths to around three minutes.

Since each of the 30 respondents were given four pairs of songs, we have 120 comparisons to analyze. In total, 55.1% correctly guessed the AI song, but by running statistical analysis, we found this was insufficient evidence to conclude that people guessed correctly a majority of the time. This tells us several things, most prominently, that AI Music has reached a level of sophistication so as to make it perceptually indistinguishable from human music in certain contexts.

We also found insufficient evidence that whether or not a person believed a certain song was AI significantly changed their average ratings of the song. Albeit, the songs believed to be

made by humans were rated higher on average. This furthers the conclusion that AI music has become functionally equivalent to human music in terms of listener experience, challenging the assumption that human-made music is impossible to replicate with AI.

This may seem to run counter to our conclusions on what AI Music is missing, but our survey also asked each participant's familiarity with music. A majority, or 55% of participants marked the lowest familiarity, "Listens to music casually", enough that we can attribute this conclusion to them. Many listeners likely could not distinguish the repetitive motifs of AI Music from human music, especially in this case of heightened ambiguity, where they didn't have access to song titles, creators, or sources.

It is important to admit, however, the limitations in these researched results. Firstly, all this data is based on purely subjective and highly personal questions about music. The fairly small sample size leaves plenty of potential for selection bias, and the survey itself was not distributed professionally.

Still, it gives us the important knowledge that, for ordinary listeners under typical conditions, the distinction between AI and human music is difficult to find. The question then becomes why and for whom human artistry still matters, something that warrants future discussion.

4. Ethical and Cultural Implications

Direct generation of music from AI has not been as discussed in the mainstream as something like AI art. AI music is more often an umbrella term including AItools that can streamline the busywork of making music. In spite of that, there have been some smaller pushes for the creation of specifically AI generated music.

On Spotify, there is a whole space of artists creating AI generated songs wholesale and putting them on the platform. Some even use AI for everything, from lyrics to titles to cover art. Needless to say, this tidal wave of low-effort music artists have been received poorly by critics. Popular album critic Anthony Fantano has been vocal in criticizing an "AI singer/songwriter" called Anna Indiana for generic and poorly made music.[5] Fantano also brings up other key points around AI in the music industry, such as the lack of transparency in its use.[6] In many cases, artists fail to communicate the extent of AI use in any given song to their audiences, leading to lots of skepticism and distrust for any music that has to do with AI. Furthering this distrust is the lack of transparency in AI itself. The vast datasets many models use for training are often complete black boxes, and whether or not songs used to train AI count as copyright infringement has been the subject of furious debate. As with most fields into which AI has expanded, this is all part of a larger struggle to protect human artists and keep the industry from being overrun by AI.

The ethical impact AI would have in music is also fiercely debated. The most pressing issue is its relationship with plagiarism, an ongoing argument across all kinds of AI. As of now, it has been decided that AI has no copyright protection, nor can it be trained from copyrighted

works. The problem lies with how vast AI training datasets can be, and with the simple fact that plagiarism through AI music cannot be caught at first glance.

Along with concerns for artistic integrity, it also raises debate over the future of the music industry. If, as our survey suggests, listeners cannot distinguish between AI and human music, will they be content with more AI on their streaming platforms? Could it take market share away from actual human creators? These are important questions that need to be answered as AI only continues to develop.

5. Conclusion

To summarize, AI music generation is fundamentally different from the process of making music naturally. The fingerprints of AI can be heard in the songs they produce, but as our survey has shown, not everyone can tell the differences in them. This raises important questions about the directions AI generated music may be taken in, and what its relationship with human music will continue to evolve into.

Works Cited

[1] Vaswani, A., & Shazeer, N. (2017). Attention Is All You Need [Review of Attention Is All You Need]. 31st Conference on Neural Information Processing Systems . https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[2] Borsos, Z., & Zeghidour, N. (2022, October 26). *AudioLM: A language modeling approach to audio generation*. Google Research. https://research.google/blog/audiolm-a-language-modeling-approach-to-audio-generation

[3] *bark#*. Bark - TTS 0.22.0 documentation. (n.d.). https://docs.coqui.ai/en/latest/models/bark.html

[4] UMG v. Suno (https://www.musicbusinessworldwide.com/files/2024/08/SUNO-response-to-copyright-suit.pdf August 1, 2024).

[5] Fantano, A. (n.d.). *This AI Song Is Terrible*. YouTube. https://www.youtube.com/watch?v=mlSlDiqngcE&ab_channel=fantano

[6] Fantano, A. (n.d.-b). *Kanye Is Beyond Washed*. YouTube. https://www.youtube.com/watch?v=JLcN9WPkYcY&t=272s&ab_channel=fantano

What Is the Most Important Part of a Chess AI? By Isaac Koshy Panicker

Abstract

Understanding the balance between depth and efficiency is crucial for developing effective AI algorithms in general. Chess AI is a challenging domain in which this balance is particularly important, as both depth and efficiency are essential for playing at a high level. This paper will examine the most efficient balance between depth and efficiency in the context of two common types of chess AI algorithms: minimax search and alpha-beta pruning. It will also discuss the impact of different components of chess AI algorithms on depth and efficiency, such as the evaluation function and the move generation algorithm. Finally, it will assess what part board evaluations and piece positions play in the efficiency and depth of an AI search. Although depth is important, it comes to a conclusion that efficiency is the leading factor that contributes the most to an algorithm's performance.

1: Introduction

This paper analyzes multiple chess playing algorithms to explain the balance between search depth and efficiency. Efficiency and depth are defined in relation to chess AI, and their importance for playing chess at a high level is discussed. The challenges of balancing depth and efficiency are presented, along with techniques that can be used to improve the efficiency of chess AI algorithms.

Artificial intelligence, or AI, uses computing to emulate human intuition and thinking processes by repeatedly training and creating algorithms. For example, the evaluation function is a function that assigns a score to each possible board position. This score is used by the chess AI algorithm to decide which move is best. The evaluation function takes into account a variety of factors, such as the material balance of the position, the safety of the king, and the potential for mobility and development.

One specific application of AI is decision making, which makes it very useful to train in board games. Chess is one of the classic implementations used by artificial intelligence, and it can be used to measure how well an algorithm can perform in the wider AI scale. Search depth and efficiency are two important factors in chess AI algorithms. Search depth is the number of moves that an algorithm can look ahead when choosing its next move, while efficiency is the speed at which an algorithm can search the game tree. Both depth and efficiency are important for playing chess at a high level, but there is a trade-off between the two. Deeper searches can lead to better moves, but they also take longer to perform.

2: Methods

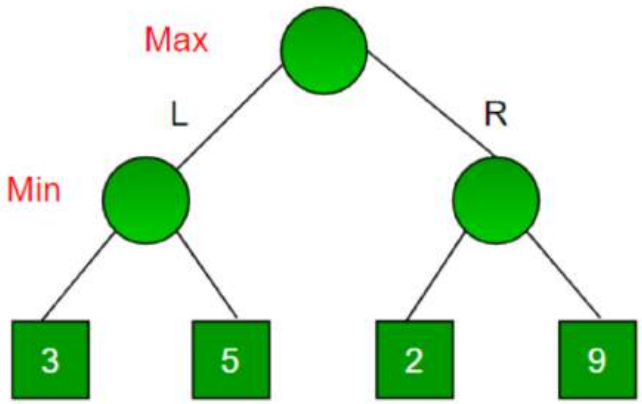
The collection of literature for this paper was conducted by searching the Internet for scholarly articles using Google Scholar and other resources. The search terms used were "chess AI," "minimax algorithm," "alpha-beta pruning," and "null move pruning." The following criteria were used to select literature sources:

The source must be a journal article or conference paper.
The source must be published within the last 10 years.
The source must be relevant to the topic of the paper, which is the balance between depth and efficiency in chess AI algorithms.

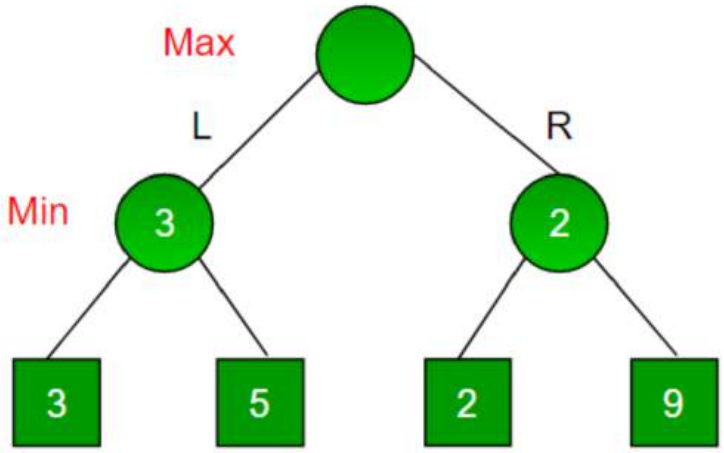
The literature review was conducted by thematically analyzing the selected sources. The identified themes included the impact of search depth and efficiency on chess AI performance, different techniques for improving the efficiency of chess AI algorithms, and the challenges of developing efficient chess AI algorithms.

3: A Standard Chess AI

One of the most common types of base search algorithms used in chess move search AI is the minimax algorithm (Aradhya, 2022). The minimax algorithm is a recursive algorithm that searches the game tree to find the best move, meaning it calls itself repeatedly. It works by considering all possible moves and countermoves for both players, and then choosing the move that leads to the best outcome for the current player. The minimax algorithm can be thought of as a two-player game. The first player is the algorithm itself (maximizer), and the second player is the opponent (minimizer). The algorithm's goal is to find the move that will lead to the best outcome, regardless of what the opponent does. From this, we can create a game tree, with paths propagating outwards. Then, it backtracks and recursively explores the game tree. At each node in the tree, the algorithm evaluates the current board position and assigns a score to it. The score represents how good the position is for the current player. To evaluate a board position, the minimax algorithm uses an evaluation function that takes into account material balance, the safety of the king, and the potential for mobility and development. Once the algorithm has evaluated all possible moves at a node, it chooses the move that leads to the highest score. This process is then repeated recursively for each of the opponent's possible moves. In the end, the minimax algorithm returns the move that leads to the highest overall score (Shannon, 1950). Here is an example:



The algorithm is focused on maximizing the AI's move score, while minimizing the score of the opposing player. This is done by making the move which leads to the highest numerical score for the AI , and the lowest score for the opponent. Because it is a backtracking algorithm, it first searches through every possible move before backtracking and choosing which move would benefit it the most. In the case of the above diagram, if the maximizer (algorithm) goes left, the minimizer would choose the move with a score of 3, because it is the least of the two options 3 and 5. If the maximizer goes right, the minimizer would choose the move with a score of 2, because it is the least of the two options 2 and 9.



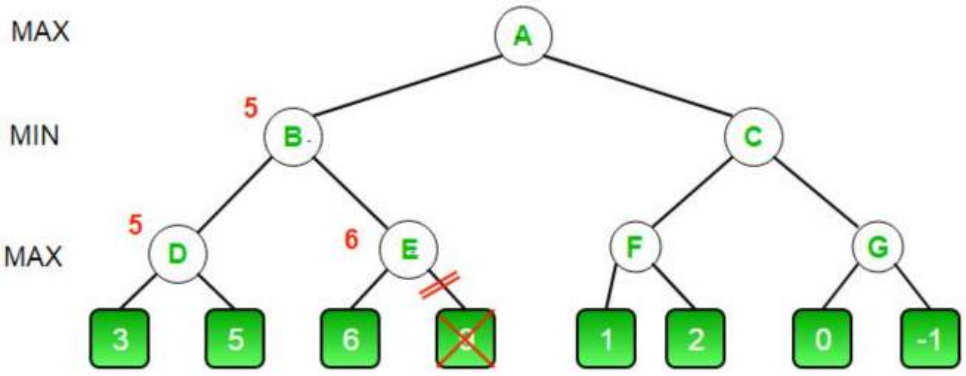
Above is the updated move tree after the AI looks ahead at future moves. The algorithm will end up going left, as it assumes the minimizer will pick 3. This is the greater of the two options so the maximizer, or AI, will choose to make the move.

4: Advanced Techniques

The minimax algorithm is a powerful tool for finding the best move in a game, but it can be computationally expensive. This is because the minimax algorithm needs to consider all possible moves and countermoves for both players. In addition to this method, many advanced chess engines such as Stockfish (Tord Romstad, Marco Costalba, and Joona Kiiski) and AlphaZero (DeepMind) use more complex techniques in order to increase their performance. To improve the efficiency of the minimax algorithm, a variety of these techniques have been developed, such as *alpha-beta pruning*.

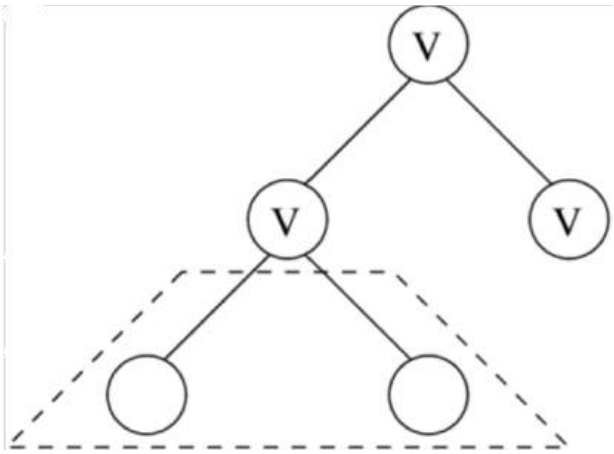
Alpha-beta pruning is one type of addition to a standard minimax algorithm that reduces the number of nodes in the move search tree. Alpha-beta pruning works by keeping track of two variables: alpha, the best possible score that the current player can achieve, and beta, the worst possible score that the current player can achieve, both assuming that the opponent plays perfectly. At each node in the game tree, the alpha-beta pruning algorithm evaluates the current

board position and assigns a score to it. If the score is greater than or equal to beta, then the algorithm can prune the entire subtree below that node. This algorithm searches through moves just as minimax would, but it stops searching down a path for a single move when it finds a possibility that the move is worse than another move. This is because the current player can never do worse than beta, regardless of what the opponent does (Shannon, 1950).



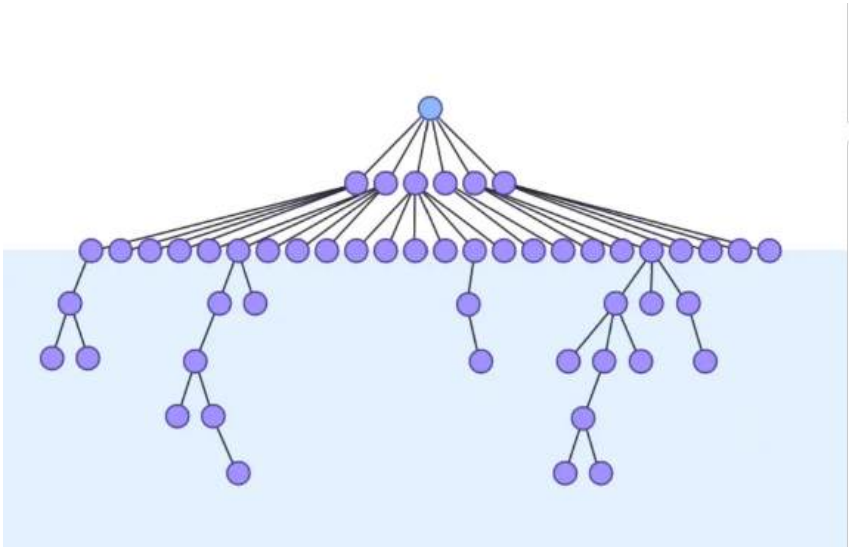
In the above diagram, from the top layer of the search tree, the maximizer searches the four layers beneath. They return that the minimizer is most likely to go left on the second layer. In the third, however, there are two options for the maximizer. It searches from move D and finds the maximum to be 5. But when searching move E, the maximizer comes across one move with a score of 6. This means it will choose move E over move D regardless of what the other branch's number is, because a move with a score of 6 is better than a move with a score of 5. The remaining branch of move E is pruned away, and doesn't have to be evaluated at all. In the end, alpha-beta returns the same final decision as minimax would, but since the engine will have to look through far fewer paths, computing power is saved and the algorithm's efficiency increases. Specifically, the average branching factor for minimax in chess originally is 35, however with the alpha-beta pruning, the program achieves a branching factor of around 25 (AI Chess Algorithms, AI Chess Project, 2025). Optimally, it would double the depth of the search tree without increasing search time. However, this is relatively more difficult to do, so move sorting, an approach where the best and worst moves are sorted, can be added to help advance the efficiency of alpha-beta.

Null move pruning is another advanced technique that can be used to further improve the efficiency of alpha-beta pruning. It works by exploiting the fact that a player can sometimes improve their position by passing their turn. Null move pruning works by considering a special move called a null move, a move that does nothing. When the null move is considered, the opponent is forced to make a move. The null move pruning algorithm then evaluates the board position after the opponent's move. If the score is greater than or equal to beta, then the algorithm can prune the entire subtree below that node.



The above diagram illustrates the pruning of a subtree on the left side. This is because the current player can never do worse than beta, regardless of what the opponent does. Null move pruning can significantly improve the efficiency of alpha-beta pruning, especially for searches that are deep enough to encounter null moves (How Does AI Play Chess? Baeldung,2023).

Quiescence search is a technique that can be used to improve the accuracy of the minimax algorithm. It works by searching the game tree deeper for positions that are considered to be quiet. Quiet positions are positions where there are no captures or checks possible. Quiescence search is used to avoid the horizon effect, a problem that occurs when the minimax algorithm is used to search to a fixed depth. The horizon effect causes the algorithm to miss winning moves that are beyond the search depth. Quiescence search can alleviate the horizon effect by searching deeper for positions that are considered to be quiet. This is because quiet positions are more likely to be the end of the game, and the best move in a quiet position is more likely to be the winning move. Below is an illustrated quiescence search (How Does AI Play Chess? Baeldung,2023).



The depth of the search tree, or number of layers is reduced by doing this. *Quiescence search* (Degni, 2023) is a further addition to alpha-beta pruning that searches for possible moves down the search tree that avoid tactical positions of the opposing side, which would put the AI at a disadvantage. The engine continues searching until its position is completely "quiet," meaning that no victorious or important moves can be made.

5: Types of Searches

Claude Shannon is considered to be the father of information theory and one of the most important computer scientists of the 20th century. In 1950, he published a seminal paper titled "Programming a Computer for Playing Chess" in which he introduced the minimax algorithm, which is still the foundation of most modern chess AI algorithms (Shannon, 1950). This paper's research on the balance between depth and efficiency in chess AI algorithms is directly related to Shannon's work on the minimax algorithm because the minimax algorithm can be computationally expensive, especially for deep searches. This research focuses on developing techniques for improving the efficiency of the minimax algorithm, and one of the most effective techniques for this is alpha-beta pruning, which was worked on by Shannon (Shannon, 1950).

According to Shannon, there are two main types of searches: brute-force and selective searches (How Does AI Play Chess? Baeldung, 2023). Brute-force searches look at every possible move that could be made with a fixed depth of layers, making them computing heavy algorithms. Selective searches specifically evaluate the best candidate moves while using more layers of the search tree, then prune away the other branches. Most programs prior to the launch of DeepMind's Alpha series in 2017 were primarily brute-force searches, with a few selective elements. More specific than the brute-force and selective search categories, there are different types of algorithms that can be implemented based on the desired output. The most basic brute-force algorithm is a depth-first search. Depth-first search is an algorithm that explores as far down as possible along each branch of the search tree before backtracking and moving on to another branch. There are many other searches for different types of games. These include Breadth-first search, which is useful in finding the shortest path in a graph and solving puzzle games, and Hill Climbing Search, used for solving computationally hard problems. Best First Search (as a path-finding algorithm,) and others.

Conclusion

This paper has discussed the balance between search depth and efficiency in chess AI algorithms. It has been shown that both depth and efficiency are important for playing chess at a high level. However, the optimal balance between depth and efficiency will depend on a number of factors, such as the available computing resources and the time available for making a move. This research has shown that advanced techniques such as alpha-beta pruning and null move pruning can be used to improve the efficiency of chess AI algorithms. However, even with these techniques, it is still difficult to search to very deep depths. This is because the branching factor of the chess game tree is very large. One way to address this challenge is to develop new search

algorithms that are more efficient than the minimax algorithm. Another way to address this challenge is to develop new hardware that is more powerful than current computers. The research has also shown that the optimal balance between depth and efficiency will vary depending on the specific chess AI algorithm being used. This is because different algorithms have different strengths and weaknesses.

Works Cited

Aradhya. “Minimax Algorithm in Game Theory.” 13 June, 2022.
“AI Chess Algorithms”, AI Chess Project
R. Degni. “The Ultimate Checkmate: AI and Chess Engines.” *Codemotion*, 20 February, 2023.
“How Does AI Play Chess?” *Baeldung*, 4 May, 2023.
E. Shannon, “Programming a computer for playing chess.” Philosophical Magazine, 1950.
Walker “The Anatomy of a Chess AI,” Medium, 20 August, 2020.

Integrative CTCF ChIP-seq and Network Analysis Reveals Promoter-Proximal Regulatory Shifts in Alzheimer's Cortex

By Shivam Wadhwa[1] and Nirupma Singh[2]

Abstract

Neurodegenerative disorders cause progressive damage to the nervous system, leading to cognitive and functional decline. Alzheimer's disease (AD) is the most common, affecting millions worldwide and lacking a cure. Understanding the molecular mechanisms driving AD is essential for developing new diagnostic and therapeutic strategies. One area that remains underexplored is the role of transcription factor binding in disease progression. This study analysed ChIP-seq data to investigate CTCF binding patterns in AD brain tissue. Data from the middle frontal area of an Alzheimer's donor was processed using R and Bioconductor packages for peak annotation and Gene Ontology (GO) enrichment. Peaks were mapped to genes, genomic locations were identified, and associated biological processes were determined. Open Targets was used to filter the gene list for AD relevance. STRING and Cytoscape were applied to build and analyse protein–protein interaction networks, and Reactome pathway analysis identified enriched pathways involving the filtered hub genes. Most CTCF binding occurred in promoter-proximal regions, with significant enrichment in processes such as small GTPase-mediated signalling, axonogenesis, and neuron projection development. Network analysis identified AKT1, GAPDH, EGFR, SRC, and IL6 as hub genes. Pathway analysis revealed enrichment in receptor tyrosine kinase signalling and PI3K/AKT pathways, both linked to neuronal survival and synaptic regulation. These findings suggest that altered CTCF binding may contribute to transcriptional dysregulation in AD. This work demonstrates the value of integrating ChIP-seq with network and pathway analysis to uncover disease-related regulatory mechanisms and potential therapeutic targets in Alzheimer's disease.

Introduction

Neurodegenerative diseases are conditions that gradually damage and destroy parts of the nervous system, particularly areas of the brain. These conditions typically develop gradually, and their effects and symptoms often appear later in life. ("Neurodegenerative Diseases") Neurodegenerative diseases occur when nerve cells in the brain or peripheral nervous system lose function over time and ultimately die. Although certain treatments may help relieve some of the physical or mental symptoms associated with neurodegenerative diseases, slowing their progression is not currently possible, and no cures exist. (Scientists) Globally, Alzheimer's affects over 55 million people and is the leading cause of dementia, while Parkinson's impacts more than 10 million, ranking second among neurodegenerative disorders. ("Statistics") Alzheimer's disease is a progressive neurodegenerative disorder characterised by memory loss, cognitive decline, and neuronal death due to abnormal protein aggregation. Its growing global

[1]Mallya Aditi International School, Bangalore, Karnataka, India.
[2]Department of Biological Sciences and Engineering, Netaji Subhas University of Technology (Formerly NSIT), Dwarka, Delhi, India.

prevalence poses a major public health challenge, straining healthcare systems and highlighting the urgent need for effective diagnostics and treatments. (Larsen et al.) However, there are several key gaps in bioinformatics research related to neurodegenerative disorders, such as a lack of high-quality, multi-modal datasets, a limited understanding of cell-type-specific mechanisms (Mathys et al.) and challenges in functional interpretation of genetic variants. (Wang et al.)

Previous studies have identified key genes such as APP, MAPT, and SNCA in Alzheimer's and Parkinson's disease, linking them to neurodegeneration pathways. GWAS and transcriptomic analyses have revealed genetic variants and dysregulated expression patterns associated with disease progression. Protein interaction studies have also uncovered critical molecular networks involved in neuronal dysfunction. (*Website*) Bioinformatics has contributed significantly to understanding neurodegenerative disorders, which is evident in studies such as discovering CSF and blood biomarkers using omics datasets. (Cheng et al.) Studies on DNA methylation, miRNAs, and lncRNAs using bioinformatics pipelines reveal regulatory changes in neurodegeneration (*Validate User*), and even the integration of GWAS data with transcriptomic and methylation profiles helps identify novel risk loci. (*Validate User*)

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is a method to identify DNA-binding regions of specific proteins or histone modifications throughout the genome. (Oba and Nakato) ChIP-Seq identifies the binding sites of DNA-associated proteins and can be used to map global binding sites for a given protein. (Hofmann-Apitius et al.) ChIP-Seq delivers genome-wide profiling with massively parallel sequencing, generating millions of counts across multiple samples for cost-effective, precise, unbiased investigation of epigenetic patterns. The reason to use Chip seq is its plethora of advantages that include its compatibility with various input DNA samples(Satoh et al.) and its ability to identify DNA targets for transcription factors or histone modifications across the entire genome of any organism. (*Chromatin Immunoprecipitation Sequencing (ChIP-Seq)*)

Even though there have been efforts to use this tool, its applications in the field of neurodegenerative disorders have been left uncharted. Hence, the clear need for further extensive research was identified, which may provide a more complete understanding of the disease. This paper analyses ChIP-seq data to identify where transcription factors bind in brain tissue from patients with neurodegenerative diseases like Alzheimer's or Parkinson's. By comparing binding patterns between diseased and healthy brain samples, the study aims to uncover regulatory changes that may contribute to disease progression. The results could help identify new genes and pathways involved in neural degeneration and suggest potential targets for therapy.

Methodology

Data Collection

A series of data studies were examined from the NCBI Gene Expression Omnibus (GEO), which included GSE261164, GSE261164, GSE231082 and GSE214911. GSE261164 described the altered PHGDH expression in astrocytes that promotes Alzheimer's pathology

using transcriptional regulation of IKKa and HMGB1, linking NFκB signalling and autophagy to disease progression. The GSE261164 study contained the ChIP-seq and scRNA-seq in brain organoids to show that astrocytic PHGDH drives Alzheimer's pathology by regulating IKKa and HMGB1, affecting NFκB signalling and autophagy. Samples of this study included healthy organoids, sporadic AD organoids, and AD organoids with PHGDH knockdown in astrocytes. The GSE214911 study had multiomics samples of postmortem DLPFC tissues from Alzheimer's and control donors, identified over 40,000 AD-specific cis-regulatory element–gene links, highlighting ZEB1 and MAFB as key transcription factors in neuronal and microglial gene regulation. Out of all the studies, the study following GSE231082 was chosen, which was related to CTCF TF ChIP-seq in the middle of the frontal area with 46 tissues of a female adult aged 90 years or above with Alzheimer's disease. This study provided the relevant differential expression and peak counts data for further analysis.

Peak Annotation and GO Enrichment

To perform the peak annotation for GSE231082, R Studio was used. Essential Bioconductor packages were loaded to facilitate the analysis, namely chipseeker, TxDb.Hsapiens.UCSC.hg38.knownGene, *org.Hs.eg.db*, and *clusterProfiler* package, which was employed for functional enrichment analysis, including Gene Ontology (GO) studies. Loading these packages ensured that all tools required for analysing gene regions from the ChIP-seq dataset and determining their potential biological roles were available. The input file, in BED format, contained the genomic coordinates of DNA regions bound by proteins, referred to as peaks. This file was read and stored as a GRanges object, enabling downstream analyses of genomic intervals. Gene model information for the human genome version hg38, including transcription start sites (TSS), exons, and introns, was retrieved from the TxDb database. This annotation resource allowed mapping of peaks to their nearest genes. Each peak was annotated with its closest gene, the corresponding genomic region (such as promoter, exon, intron, or intergenic), and its associated Entrez gene identifier. Peaks within ±3000 base pairs of the TSS were classified as promoter-proximal. Additional biological metadata, such as gene symbols, were incorporated from the *org.Hs.eg.db* database. The distribution of peaks across different genomic regions was visualised using a pie chart, while the distances of peaks to the nearest TSS were represented through a histogram.

The annotated data were then converted to a tabular format, and the list of associated gene identifiers was extracted after removing unmapped entries. Gene Ontology enrichment analysis was performed using the "Biological Process" (BP) ontology to identify significantly overrepresented processes among the genes associated with ChIP-seq peaks. A p-value cutoff of 0.01 was applied to ensure statistical significance. This analysis revealed biological processes in which these genes were predominantly involved, including those related to cellular regulation, neural function, and disease-associated pathways. The top ten enriched GO Biological Processes were visualised as a bar chart to provide a clear summary of functional categories represented in the dataset.

Network and Pathway analysis

The comprehensive list of genes, taken from the Chip seq analysis, is pasted in the Open Targets database, a database that links genes to diseases using multiple evidence sources (genomics, transcriptomics, genetics. This is done to see which genes are strongly associated with Alzheimer's. By keeping only the genes with strong evidence scores, this shortened the list by removing genes with little or no known relevance. After filtering the gene list using Open Targets, the new gene list was entered into the STRING-Search Tool for the Retrieval of Interacting Genes/Proteins, a resource for retrieving and visualising known and predicted protein–protein interactions. The multiple proteins option was selected, the gene symbols were pasted into the search bar, *Homo sapiens* was chosen as the organism, and the analysis was initiated. The resulting interaction network was examined, and in the case of a fully connected network, the confidence setting was adjusted to high confidence. The network was then exported both as a high-resolution PNG image and as a short tabular text output (TSV) listing the one-way edges (A–B). The TSV file was opened in Excel/Google Sheets, and the "node 1" and "node 2" columns were extracted and saved as a CSV file. This file was imported into Cytoscape, a software platform for visualising, analysing, and integrating complex networks.

The network was analysed using Cytoscape's built-in network analysis tool, with the directed network option left unchecked. The resulting topological parameters table was exported and examined, focusing on the columns for degree, betweenness centrality, closeness centrality, clustering coefficient, and name. The table was sorted by degree from highest to lowest, and the top ten genes with the highest degree values were identified. The complete gene list from the parameter file was entered into the Reactome database, a curated knowledgebase of biological pathways and reactions in human biology. The analysis tool was selected, the genes were pasted into the search bar, and the analysis was run with default settings. The results page displayed a network of pathways along with a table of top pathways ranked by relevance. From the download menu associated with the table, all five available output files were downloaded. The results file was examined to determine whether the top ten genes identified in the previous network analysis were present within the top 25 pathways. Each of the top ten genes was searched within the file, and if a gene was found in the top 25 pathways, it was marked as present; otherwise, it was noted as absent.

Results

Peak Annotation analysis

After performing the peak annotation analysis, the top annotated genes were identified. They were namely: chloride voltage-gated channel 4 (CLCN4), glial cells missing transcription factor 1 (GCM1), regulating synaptic membrane exocytosis 1 (RIMS1), RNA binding motif protein 17 (RBM17), keratin 42 (KRT42P), pseudogene (POTEM), long intergenic non-protein coding RNA 3007 (LINC03007), and neutrophil cytosolic factor 4 (NCF4). The pie chart (Figure 1) shows the distribution of categories that represent the proportion of peaks mapped to specific

genomic features**:** The largest fraction of peaks is found in Promoter regions (<=1kb). Intronic regions (1st Intron + Other Intron) together account for ~37%. Distal intergenic regions are 16.36% and very few peaks are found in UTRs or downstream regions. In the stacked bar plot(Figure 2), the width of each colored band shows how many binding sites fall within that distance category. Most transcription factor-binding sites are concentrated close to the Translation Start Site (TSS) (0–1 kb range), indicating that promoters are a major binding region. Binding density tapers off gradually as distance increases from the TSS, with fewer sites in regions >100 kb away.

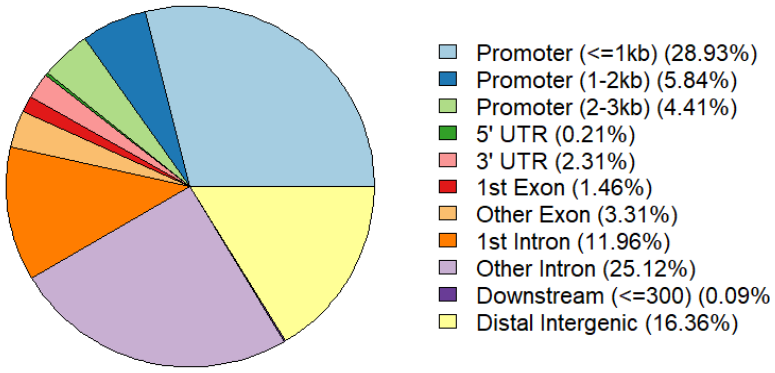


Figure 1. A pie chart showing the distribution of peaks in the GSE. Each colour represents a particular category as follows: Light Blue – Promoter (≤1 kb) – 28.93%, Dark Blue – Promoter (1–2 kb) – 5.84%, Light Green – Promoter (2–3 kb) – 4.41%, Dark Green – 5′ UTR – 0.21%, Red – 3′ UTR – 2.31%, Peach/Beige – 1st Exon – 1.46%, Orange – Other Exon – 3.31%, Bright Orange – 1st Intron – 11.96%, Lavender/Purple – Other Intron – 25.12%, Dark Purple – Downstream (≤300 bp) – 0.09% Light Yellow – Distal Intergenic – 16.36%

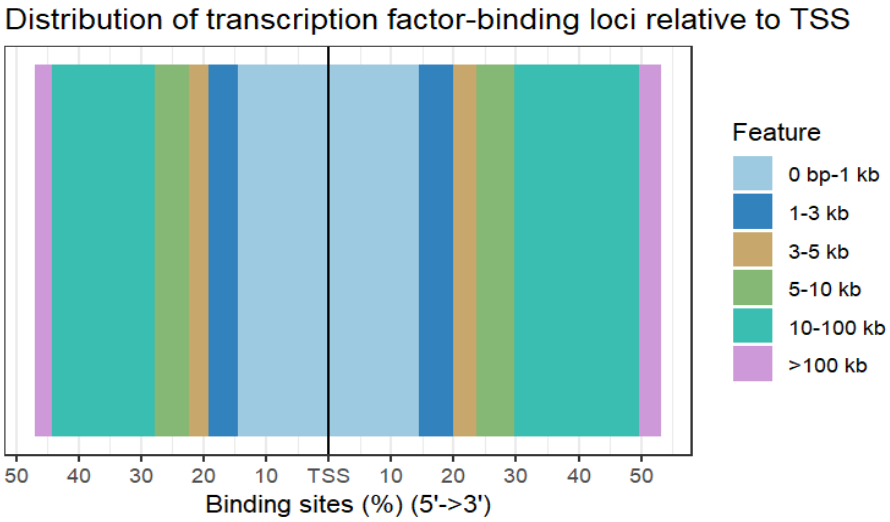


Figure 2. A stacked bar plot showing the distribution of transcription factor-binding sites relative to TSS. The colored bands represent distance categories from the TSS: Light Blue – 0 bp–1 kb, Dark Blue – 1–3 kb, Brown – 3–5 kb, Green – 5–10 kb, Teal – 10–100 kb, Purple – >100 kb

GO Enrichment Analysis

The gene ontology analysis (Figure 3) of GSE231082 showed the different types of biological processes associated. They were: small GTPase-mediated signal transduction, axonogenesis, regulation of neuron projection development and positive regulation of cell projection organisation.

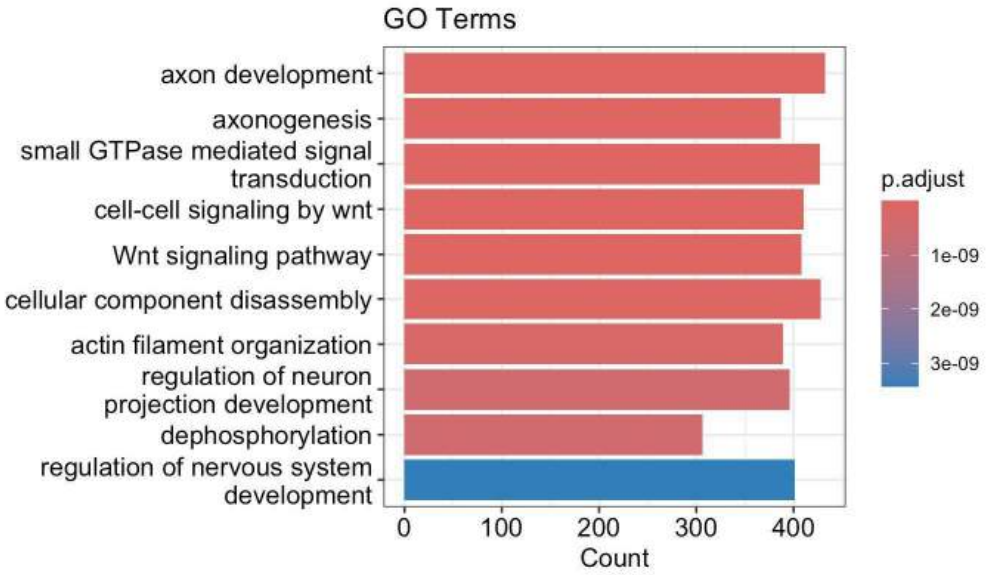


Figure 3. A bar plot displays the top enriched Gene Ontology Biological Processes among genes associated with ChIP-seq peaks. Colour corresponds to the adjusted p-value of enrichment: redder bars denote more statistically significant enrichment (smaller adjusted p-values, ~3×10^{-10}), whereas bluer bars indicate relatively less significant terms (up to ~9×10^{-10}).

Network analysis

From an initial set of 89,000 genes, Open Targets filtering with a high global score reduced the list to 1326 Alzheimer associated genes. STRING helped in providing the interactions for the final list of genes. In Cytoscape, a network of 1326 nodes and 19,260 edges was visualised. Four topological parameters were computed for each node. The degree measured the number of direct connections a node had, identifying hub genes. Betweenness centrality quantified how often a node lay on shortest paths between other nodes, indicating its role as a connector. Closeness centrality measures a node's average distance to all others, reflecting its accessibility within the network. The clustering coefficient assessed the tendency of a node's neighbours to be interconnected, indicating local modularity. Nodes were ranked by degree to identify the top hub genes for downstream pathway enrichment and interpretation.

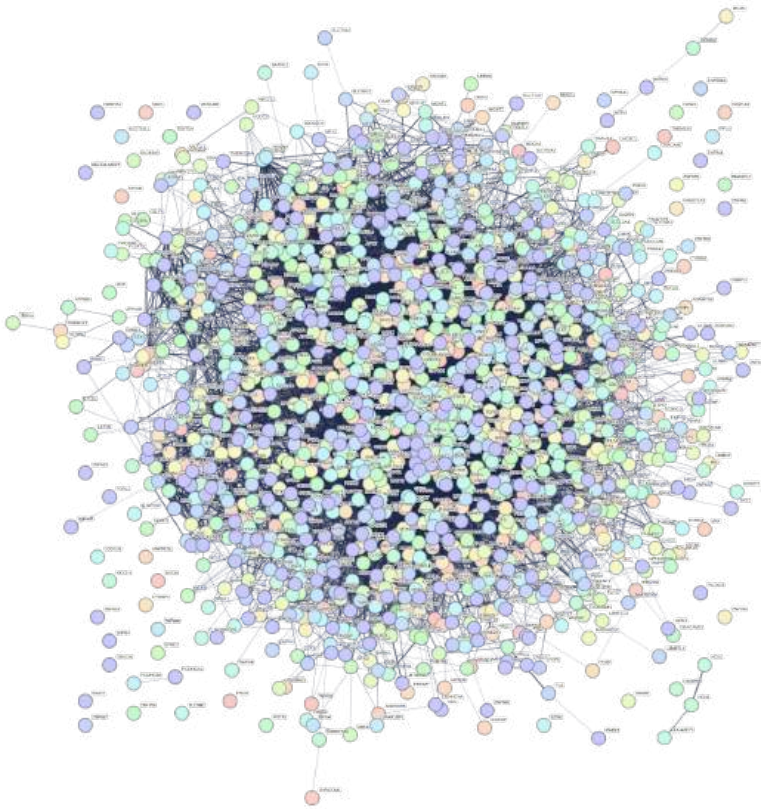


Figure 4. A figure showing the protein–protein interaction network generated from the gene list. Each node represents a protein encoded by a gene from the filtered ChIP-seq dataset. Node colours correspond to distinct biological processes.

S.No.	Name	Degree	Betweenness Centrality	Closeness Centrality	Clustering Coefficient
1	AKT1	340	0.04879645425	0.5344129555	0.1481693562
2	GAPDH	317	0.0512709902	0.53182917	0.1670127381
3	EGFR	270	0.03548684748	0.5069124424	0.1708660333
4	SRC	264	0.03447401677	0.5065234075	0.1702961171
5	IL6	246	0.01462524214	0.4938271605	0.2267794923
6	ALB	231	0.0149815669	0.4956815621	0.2346320346
7	IL1B	229	0.01010119854	0.4876246768	0.2432774075
8	JUN	227	0.01690979595	0.4956815621	0.2297766169
9	STAT3	226	0.01254918073	0.4892512973	0.2462536873
10	PTEN	211	0.02174621749	0.4979253112	0.2184608441

Table 1. Network Topological Parameters for Key Genes

Pathway analysis

After performing pathway analysis for the final list of genes, the top pathways that were identified were - Signaling by Receptor Tyrosine Kinases (R-HSA-9006934), Diseases of signal transduction by growth factor receptors and second messengers(R-HSA-5663202), Signal Transduction(R-HSA-162582), Negative regulation of the PI3K/AKT network(R-HSA-199418) and PI5P, PP2A and IER3 Regulate PI3K/AKT Signaling (R-HSA-6811558). From the final gene list, the top ten genes were chosen. A search function was performed, and each of the genes was searched to check if they were among the top 25 pathways. Out of the finalised 10 genes, 6 were found to be present in the top 10 pathways. They were namely - AKT1, EGFR, SRC, JUN, STAT3 and PTEN.

Discussion

This study used ChIP-seq data to examine CTCF binding in Alzheimer's disease brain tissue. Most peaks were in promoter regions, with many within 1 kb of the TSS, showing a preference for proximal regulatory sites. GO enrichment linked the genes to small GTPase signalling, axonogenesis, and neuron projection development, all relevant to AD. Open Targets filtering focused the list on Alzheimer associated genes. STRING and Cytoscape network analysis highlighted AKT1, GAPDH, EGFR, SRC, and IL6 as key hubs. Reactome pathway analysis showed enrichment in receptor tyrosine kinase and PI3K/AKT signalling. These results suggest that altered CTCF binding may affect critical signalling and neuronal processes, providing insight into potential molecular mechanisms in AD progression.

Previous work has profiled transcriptional and chromatin states in Alzheimer's tissue, but seldom mapped transcription factor binding genome-wide in the patient's brain. Single-nucleus multiomics studies combined snRNA-seq with snATAC-seq to link cell-type expression changes to accessible regulatory elements and nominated regulators such as ZEB1 and MAFB, yet they did not directly provide TF occupancy maps by ChIP-seq.(Anderson et al.) Another example is the Meta-analyses of bulk and single-cell transcriptomes characterised widespread differential expression and produced consensus disease signatures, but these approaches infer regulation indirectly rather than measuring protein–DNA binding.(Wan et al.) Similar studies include - In the study of diabetic nephropathy, genes differentially expressed in kidney tissue (from microarray datasets GSE30528 and GSE33744) were analysed via GO and pathway enrichment, along with Ingenuity Pathway Analysis (IPA) to identify hub genes and pathways involved in disease progression (Geng et al.)

This research relied on ChIP-seq datasets that are publicly available, which in turn constrained the sample size and variety. The study focused on the CTCF transcription factor and the middle frontal area 46 of the brain. Therefore, the study's outcomes are unlikely to reflect the full range of transcriptional regulatory changes that occur in other brain regions, other transcription factors, or in the context of Alzheimer's disease. Also, the analysis lacked experimental validation of the predicted interactions, pathways, or gene functions, and all the work was done in silico. This presents a risk of biological inaccuracies and misinterpretation of

underlying mechanisms. Dependence on a single type of data is another constraint. Although gene regulation data captured through ChIP-seq are useful, it does not capture changes in gene expression, chromatin level, or proteins quantitatively. Without the inclusion of complementary datasets like RNA-seq, ATAC-seq, or proteomic analysis, it remains unclear whether the changes in binding in question are impactful changes at the gene or protein level. Also, since existing pathway and interaction databases were used, the analysis is limited to the current biological frameworks and may ignore novel or less characterised interactions.

In conclusion, this study demonstrated the use of ChIP-seq analysis to investigate CTCF binding patterns in Alzheimer's disease brain tissue. The results showed a strong preference for promoter-proximal binding and identified genes linked to processes such as small GTPase signalling, axonogenesis, and neuron projection development. Filtering through Open Targets focused the analysis on Alzheimer associated genes, while STRING and Cytoscape network analysis revealed key hub genes including AKT1, GAPDH, EGFR, SRC, and IL6. Reactome pathway enrichment highlighted receptor tyrosine kinase and PI3K/AKT signalling as central pathways. These findings suggest that altered CTCF binding may contribute to disease-related regulatory changes, offering potential targets for further investigation and advancing understanding of Alzheimer's disease mechanisms.

Acknowledgements

The authors would like to acknowledge Aashna Saraf, Founder of CreatED, for providing valuable feedback and guidance throughout the project.

Works Cited

Anderson, Ashlyn G., et al. “Single Nucleus Multiomics Identifies ZEB1 and MAFB as Candidate Regulators of Alzheimer’s Disease-Specific Cis-Regulatory Elements.” Cell Genomics, vol. 3, no. 3, Feb. 2023, p. 100263.

Cheng, Liang, et al. Omics Data Integration towards Mining of Phenotype Specific Biomarkers in Cancers and Diseases. Frontiers Media SA, 2022.

Chromatin Immunoprecipitation Sequencing (ChIP-Seq). https://sapac.illumina.com/techniques/sequencing/dna-sequencing/chip-seq.html. Accessed 26 July 2025.

Geng, Xiao‐dong, et al. “Identification of Key Genes and Pathways in Diabetic Nephropathy by Bioinformatics Analysis.” Journal of Diabetes Investigation, vol. 10, no. 4, Jan. 2019, p. 972.

Hofmann-Apitius, M., et al. “Bioinformatics Mining and Modeling Methods for the Identification of Disease Mechanisms in Neurodegenerative Disorders.” International Journal of Molecular Sciences, vol. 16, no. 12, Dec. 2015, https://doi.org/10.3390/ijms161226148.

Larsen, Frode Kibsgaard, et al. “Age of Alzheimer’s Disease Diagnosis in People with Down Syndrome and Associated Factors: Results from the Horizon 21 European Down Syndrome Consortium.” Alzheimer’s & Dementia, vol. 20, no. 5, May 2024, pp. 3270–80.

Mathys, Hansruedi, et al. “Single-Cell Transcriptomic Analysis of Alzheimer’s Disease.” Nature, vol. 570, no. 7761, May 2019, pp. 332–37.

“Neurodegenerative Diseases.” Cleveland Clinic, https://my.clevelandclinic.org/health/diseases/24976-neurodegenerative-diseases. Accessed 24 July 2025.

Oba, Gina Miku, and Ryuichiro Nakato. “A Beginner’s Guide to ChIP-Seq Analysis.” The Biochemist, vol. 47, no. 1, Mar. 2025, pp. 21–26.

Satoh, J., et al. “Pathway Analysis of ChIP-Seq-Based NRF1 Target Genes Suggests a Logical Hypothesis of Their Involvement in the Pathogenesis of Neurodegenerative Diseases.” Gene Regulation and Systems Biology, vol. 7, Nov. 2013, https://doi.org/10.4137/GRSB.S13204.

Scientists, All. “Neurodegenerative Diseases.” National Institute of Environmental Health Sciences, https://www.niehs.nih.gov/research/supported/health/neurodegenerative. Accessed 24 July 2025.

“Statistics.” Parkinson’s Foundation, https://www.parkinson.org/understanding-parkinsons/statistics. Accessed 26 July 2025.

Validate User. https://jamanetwork.com/journals/jamaneurology/fullarticle/1150044. Accessed 26 July 2025.

---. https://academic.oup.com/nar/article/49/D1/D605/6006194?login=false. Accessed 26 July 2025.

Wang, Minghui, et al. “The Mount Sinai Cohort of Large-Scale Genomic, Transcriptomic and Proteomic Data in Alzheimer’s Disease.” Scientific Data, vol. 5, no. 1, Sept. 2018, pp. 1–16.

Wan, Ying-Wooi, et al. “Meta-Analysis of the Alzheimer’s Disease Human Brain Transcriptome and Functional Dissection in Mouse Models.” Cell Reports, vol. 32, no. 2, July 2020, p. 107908.

Website. https://doi.org/10.1016/j.cell.2013.03.030.

Emerging AI Frontiers Against Digital Fraud: A Global Perspective on Next-Gen Fraud Prevention Tools By Czar Kunder

Abstract

This research examines the potential of new artificial intelligence (AI) technologies in tackling the world-wide fight against digital fraud. While many existing fraud detection technologies currently leverage **machine learning (ML)** and **natural language processing (NLP)** to assist in identity verification and detection of fraudulent activities, these technological advancements lack other dimensions of flexibility, data protection, and scalability to navigate next developments in fraud. As a consequence, this study suggests new next generation AI technologies such as **Graph Neural Networks (GNNs)**, **Federated Learning**, and **Reinforcement Learning**, which have potential applicability in aiding the detection of rapid advances in cross-boundary fraud. The study employed a mixed methods framework by incorporating a systematic literature review with real-world primary data collected from an online digital survey with a diverse sample group. The research also proposes how financial literacy can impact public familiarity, trust and level of engagement applied to emerging public and private sector fraud detection technologies. This framework overall aims to provide future-focused, ethical, and scalable approaches to fraud prevention, and thereby brings a new conceptual viewpoint to future discussions and research on cybersecurity.

1. Introduction

Digital fraud presents an increasing challenge in a more digitized world. From phishing and identity theft to false e-commerce sites and transaction laundering, cybercriminals are exploiting technology limitations and social engineering techniques, at an alarming rate. The Federal Trade Commission (2023) reported that the total global financial losses to online fraud exceeded $ 8.8 billion in 2022, a whopping 44% increase from the previous year. With these figures, it is clear that more intelligent, adaptable and rigorous fraud detection systems are required.

In recent times, AI has been considered an essential weapon in combating digital crime. Traditional AI technologies such as supervised machine learning, data mining, and NLP make it possible for institutions to monitor, detect, and take action against anomalous financial activities based on past patterns. However, the capabilities for identifying and preventing digital crime are becoming increasingly limited by their dependence on external data based on pre-defined datasets, inability to preserve user privacy, and difficulty in identifying fraud rings or new fraud schemes (Ngai et al., 2011; West & Bhattacharya, 2016).

Recent developments in AI research offer some exciting new possibilities. Graph Neural Networks, for instance, can model complex transactional relationships, Federated Learning allows for decentralized model training in a privacy-preserving manner, and Reinforcement Learning has the ability to respond and adapt to continuously evolving fraud tactics. (Zhao et al.,

2021). Even with these promising technologies we know little about their abilities, especially within scalable real-time global applications.

Summary of Existing Literature

Prior studies have noted AI's utility in fraud detection, especially in relation to banking and e-commerce. While there has been significant research about the classical models, little attention has been given to new paradigms in neural architectures or on ethical issues surrounding data sharing and access to data. Additionally, there has been a call for interdisciplinary solutions that promote that are technically advanced but are also formulated in socially, ethically, and policy-aware frameworks (Bhattacharyya et al., 2011; Abbas et al., 2021).

Research Questions

- How might emerging artificial intelligence technologies—namely, Graph Neural Networks, Federated Learning and Reinforcement Learning—change the way we combat digital fraud globally?
- What are the limitations of traditional AI models such as machine learning and natural language processing for detecting ever-evolving patterns of digital fraud?
- What are the roles of public awareness, trust in AI, and financial literacy in enabling the optimum effectiveness and utilization of AI-based fraud prevention tools?
- What ethical, privacy and regulatory considerations come into play in the implementation of deep AI solutions for fraud detection around different global jurisdictions?

Objectives of the Study

- To evaluate the technical and ethical strengths of emerging AI tools in fraud prevention.
- To explore the boundaries of standard AI systems encountering new forms of digital fraud.
- To assess public sentiment, awareness, trust, and financial literacy in relation to AI-based fraud prevention systems.
- To propose a hybrid AI framework that integrates emerging and conventional techniques for scalable implementation.

2. Literature Review

Conventional AI Techniques in Fraud Detection

The literature shows that ML algorithms are commonly used and tested in financial fraud detection. Decision trees, logistic regression, and support vector machines have good success in pattern recognition and classification, but do not perform well against adversarial manipulation and imbalance in fraud datasets. NLP is often used to assess customer reviews, and phishing emails, but it is greatly limited by the variability of language and context details (West, & Bhattacharya, 2016).

Anomaly detection, another commonly used technique, focuses on flagging outlier behavior. While computationally efficient, it lacks contextual depth and interpretability—two features increasingly important for compliance and user trust.

Limitations of Existing Models

Traditional methods depend heavily on labeled historical data, which can quickly become outdated. These systems are also reactive rather than predictive, failing to adapt swiftly to novel scams. Moreover, concerns about surveillance, data ownership, and AI bias have led to calls for more transparent and privacy-conscious models (Ngai et al., 2011; Bhattacharyya et al., 2011). Traditional techniques rely on labeled historical data which quickly become outdated and are also inherently reactive rather than predictive as they are not designed to keep up as fast to new scams. Additionally, fears of surveillance, data ownership, and AI bias has led to a call for more transparent and privacy-sensitive models (Ngai et al., 2011; Bhattacharyya et al., 2011).

Emerging AI Solutions.

New literature illustrates the potential of advanced tools like GNNs, utilized for encoding the relational structure of fraud networks, in identifying hidden links of entities that rules-based systems may fail to do. Federated Learning provides a privacy-friendly model that allows institutions to work together to identify fraud without needing to share user data. Lastly, Reinforcement Learning brings a feedback mechanism allowing the model to learn as tactics adjust dynamically, which is a big benefit compared to the more static nature of existing models (Chen et al., 2023; Zhao et al., 2021).

The Role of Financial Literacy in AI-Driven Fraud Prevention

Beyond the capabilities of AI systems, user education and financial literacy are important in determining how users interact with, and trust fraud prevention technologies. Users who are more digitally financially literate are also better at recognizing fraud, and are more inclined to use fraud prevention tools, as research suggests (Lusardi & Mitchell, 2014; OECD, 2021). Including financial literacy as part of the fraud defense will not only help users become more resilient but will enhance the efficacy of AI systems.

3. Data Analysis and Procedures

3.1 Research Design

A **cross-sectional online survey** was conducted to explore individuals' experiences with digital scams, their preventive behaviors, their opinions on the role of AI in fraud detection, factors of trust and concern, willingness to adopt AI tools, and perceptions regarding technologies, regions, and barriers to implementation.

3.2 Sampling and Recruitment

A **non-probability convenience/voluntary response** sampling approach was used. The Google Form link was distributed through online channels (e.g., personal networks and social platforms); respondents self-selected into the study. No demographic screening questions were included, by design, so analysis is **unstratified** and focuses on behaviors and perceptions rather than age/education/employment subgroups. The final analytic sample comprised **N = 39 complete responses**.

A convenience/voluntary response sampling approach was used that is not reliant on a **probability-based methodology.** The Google Form link was disseminated through online (e.g., personal networks and social platforms) environments where respondents self-selected into the study. The online survey was designed with no demographic screening questions, meaning the analysis was **not stratified** and the focus was behaviors and perceptions rather than subgroups of age/education/employment. **N = 39 complete responses** were retained for the final analytic sample.

Limitations: Due to the sample not being random, findings cannot be statistically generalizable to the larger population; findings may also reflect selection bias towards respondents who are internet-active and self-reported experiences.

3.3 Instrument

The questionnaire contained close-ended items organized into the following domains, mirroring the figures displayed in Google Forms:

1. **Digital scam exposure and types**
    - Encountered a digital scam in the last 5 years (Yes/No).
    - Types experienced (multiple selection): e.g., phishing (email/SMS), OTP/UPI/bank fraud, fake/clone sites, investment scams, social-media impersonation, tech-support scams, online-shopping fraud, other, or none.
2. **Preventive behaviors and information sources**
    - Frequency of verifying authenticity of financial/personal requests (Always/Often/Sometimes/Rarely/Never).
    - Primary sources for digital-safety information (multiple selection): government portals, corporate websites, social media, friends/family, news media, school/college, etc.
3. **Perceptions and awareness of AI in fraud detection**
    - Perceived effectiveness of AI vs. traditional methods (5-point Likert from Strongly disagree to Strongly agree).
    - Awareness that banks/apps use AI for fraud detection (Yes/Somewhat/No).
    - Belief that AI can better adapt to Indian scam patterns (Yes/Maybe/No).
4. **Trust drivers and concerns**
    - Factors that would increase trust (multiple selection): transparency/explain ability, government regulation/approval, positive user reviews, real-time alerts, none.

    - Concerns (multiple selection): privacy, false alarms, lack of human support, data misuse, none.
5. **Adoption and usage of AI tools**
    - Preference for financial apps with built-in AI scam alerts (Definitely/Maybe/Not really/Never).
    - Willingness to use an AI tool to scan suspicious messages/emails (Yes/Maybe/No).
    - Prior use of AI-based customer service (Yes—helpful; Yes—not helpful; No; Not sure).
    - Support for a nationwide AI-based anti-scam system (1–5 Likert).
6. **Technologies, regions, and implementation challenges**
    - Regions perceived as leading in AI-based fraud prevention (single choice).
    - Most promising AI technologies (multiple selection): Machine Learning, NLP, Computer Vision, Behavioral Biometrics, Predictive Analytics, "Not familiar," others.
    - Biggest global challenges (select up to three): data-privacy concerns, lack of skilled talent, cost of AI, regulatory restrictions, algorithmic bias, lack of awareness/trust, others.

3.4 Data Preparation

Data was transferred from Google Forms to **Google Sheets/Microsoft Excel,** eliminating partial submissions (N=39). Responses were reviewed for overt duplicates and internal inconsistencies. For multiple selection items, each selection was b**inary-coded** (1= selected; 0= not selected), acknowledging that column totals may be greater than 100% by design. Additionally, Likert items were **numerically coded** (e.g., 1-5).

3.5 Statistical Analysis

Analysis primarily used **descriptive statistics**:

- **Frequencies and percentages** for all categorical items.
- For Likert-scale items, **means, medians, standard deviations, and quartiles** were computed to characterize central tendency and dispersion.
- **Cross-tabulations** were defined ex ante for core behavioral relationships using variables situated in the instrument (serious scam-exposure [Yes/No] × verification frequency; awareness of AI use × adoption willingness; trust factors × support for a national system). Where insightful, row-percentage heatmaps / bar charts were produced from the crosstabs for visualizing differences.

All figures in the Results section were generated from Google Forms/Sheets (pie charts, bar charts, and Likert distributions), with sample size **N = 39** reported throughout.

3.6 Ethical Considerations

Participation in the research was entirely voluntary. The survey began with a short introductory statement outlining the purpose and process. No personal-identifying information was gathered; all responses are anonymous, and all responses will be stored securely and used strictly for research purposes only.

4. Results

This section presents the results from the survey undertaken to investigate public exposure to and attitudes toward digital scams, as well as the trust and knowledge barriers towards using AI-enabled fraud detection systems.

4.1 Digital Scam Exposure and Types

A considerable **69.2%** of respondents claimed they encountered a digital scam within the last five years, which indicates a high level of vulnerability including people who claim to be digitally literate. Therefore, only a small minority, just **30.8%**, could honestly claim to have never experienced some form of digital fraud which supports the conclusion that the threat of online scams is pervasive and has become deeply embedded in the current digital ecosystem. The order of the types of scams commonly experienced by respondents were **phishing (41%)**, delivered via deceitful emails or SMSs, followed by **social media impersonation (30.8%)** and **online shopping scams (25.6%)**. Both social media and digital commerce environments are common presentable avenues for deception. **Tech support scams (23.1%), investment fraud (17.9%)**, and **UPI/bank fraud (17.9%)** were also weighed heavily, alluding to the sophistication and variability of digital deception.

Fig 1.1

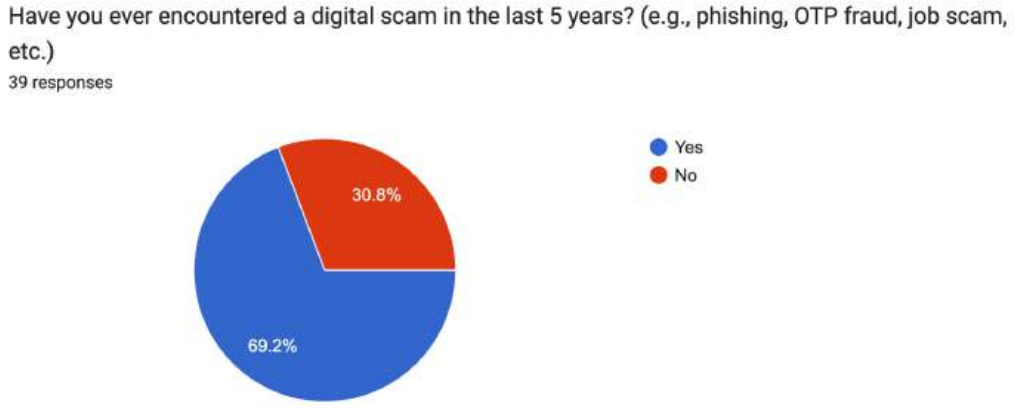
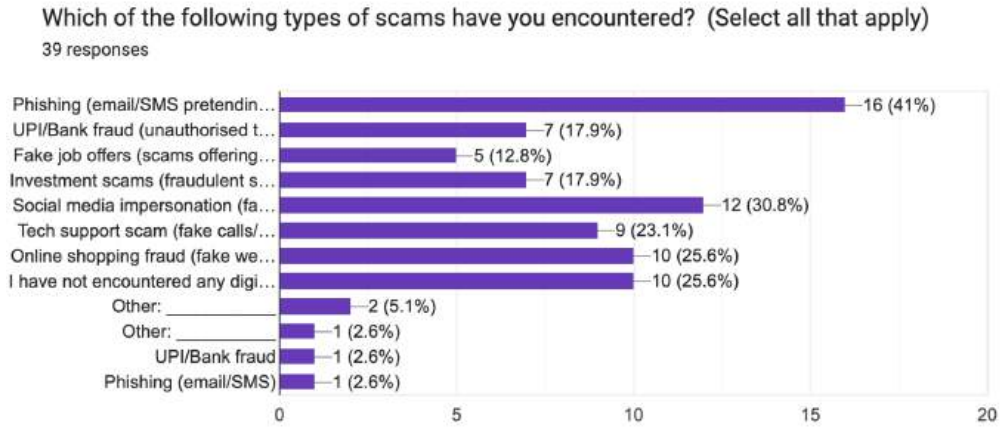


Fig 2.2



4.2 Verification Habits and Information Sources:

Despite the high frequency of respondents exposed to scams, only **25.6%** of the participants in this study indicated they "always" verify whether messages or requests are suspicious. More participants, **35.9%**, said they verify the authenticity of messages or requests "often," while the other participants surveyed said they verify the authenticity of messages or requests "sometimes" or "rarely." The findings indicate that being attentive to scams seems to exist, however habitual verification of messages or requests is still developing among scam aware individuals.

When asked where they typically learned about fraud prevention, respondents identified **social media(43.6%)** and **news media (35.9%)** as their main sources of information. Only **7.7%** of respondents referred to government portals as their source, and **2.6%** said educational institutions had a role, indicating little involvement by formal institutions in sharing knowledge about digital safety. Their informal knowledge reliance could contribute to varying and possibly out-of-date understanding of changing fraud tactics.

Fig 1.3

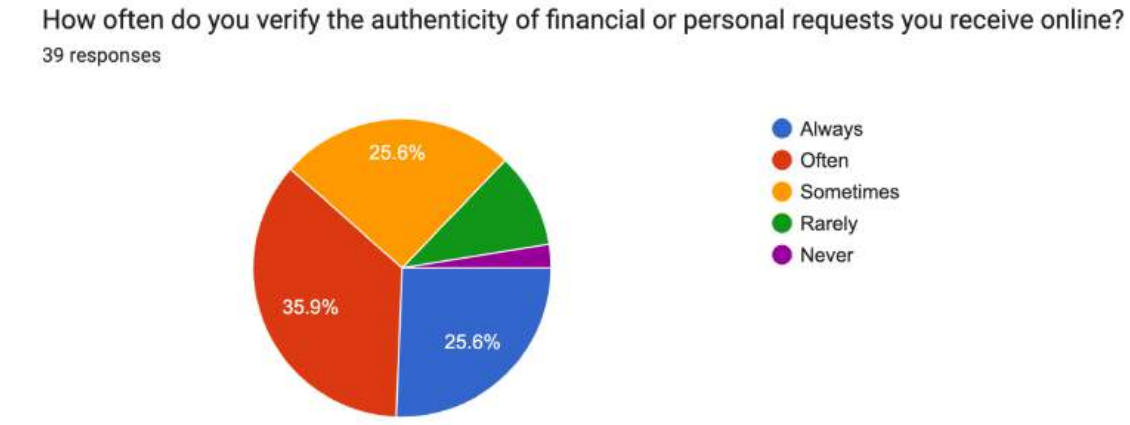
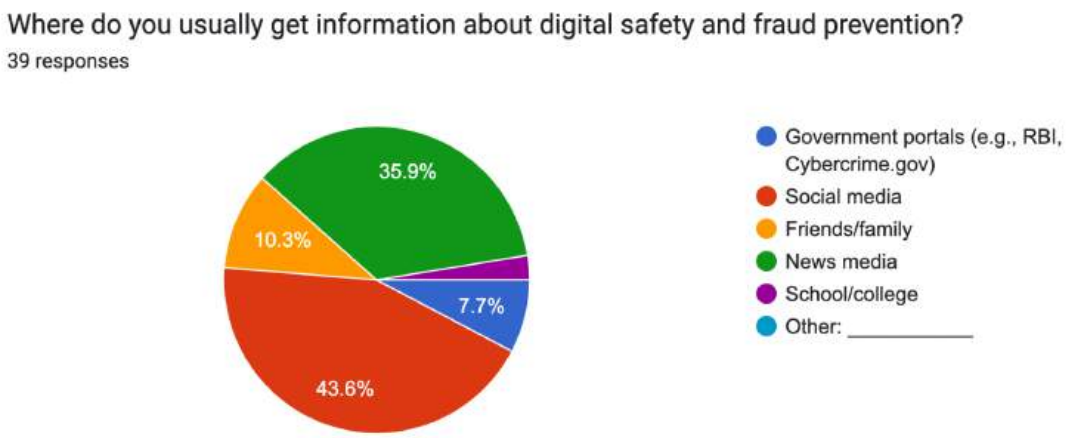


Fig 1.4



4.3 Perceptions of AI in Fraud Detection

Participants expressed moderate confidence in AI's effectiveness. While **38.5%** agreed and **17.9% strongly agreed** that AI is more effective than traditional fraud detection systems, a sizable **38.5%** remained neutral. Only **5.1%** expressed outright disagreement. Notably, only **43.6%** were aware that AI is currently used in banking and fintech fraud detection tools—suggesting a critical gap between **institutional adoption** and **public awareness**.

Participants had moderate levels of confidence in AI's capabilities. For example, while **38.5%** agreed and **17.9%** strongly agreed that AI is more efficient than traditional fraud detection systems, a significant number, namely **38.5%**, were neutral. Only 5.1% of respondents strongly disagreed. Also importantly, only **43.6%** knew that AI is currently used in banking and fintech fraud detection tools indicating a major gap between **institutional adoption** and **public awareness.**

When asked if AI would adapt better to Indian scam patterns that have a mix of language variations, cultural signals, and local fraud structures, **56.4%** answered "Yes" and **33.3%** stated "Maybe". This highlights a large amount of public optimism with regards to AI being contextually adaptive in the Indian space.

Fig 1.5

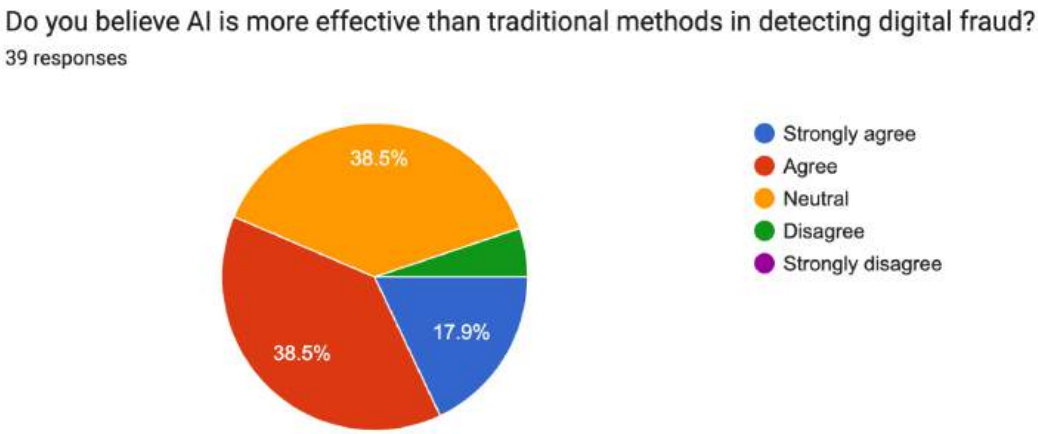


Fig 1.6

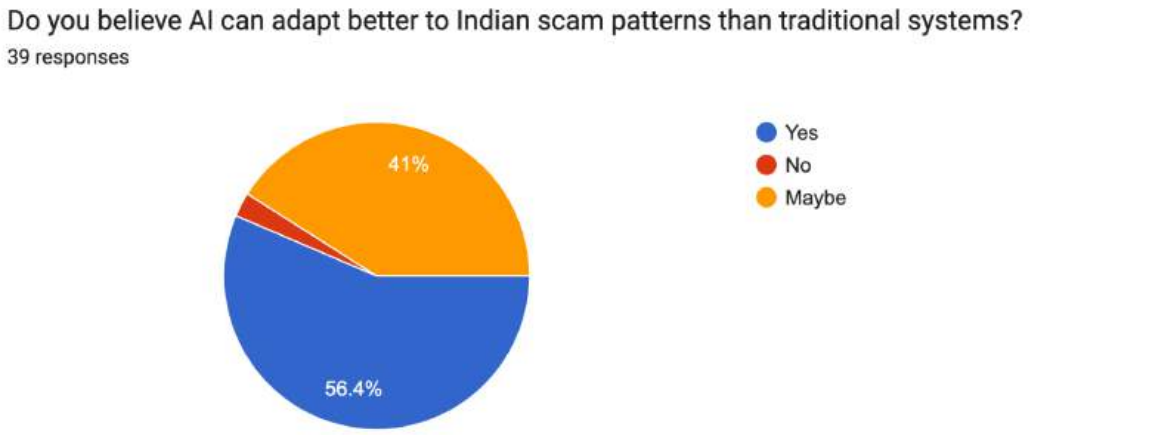
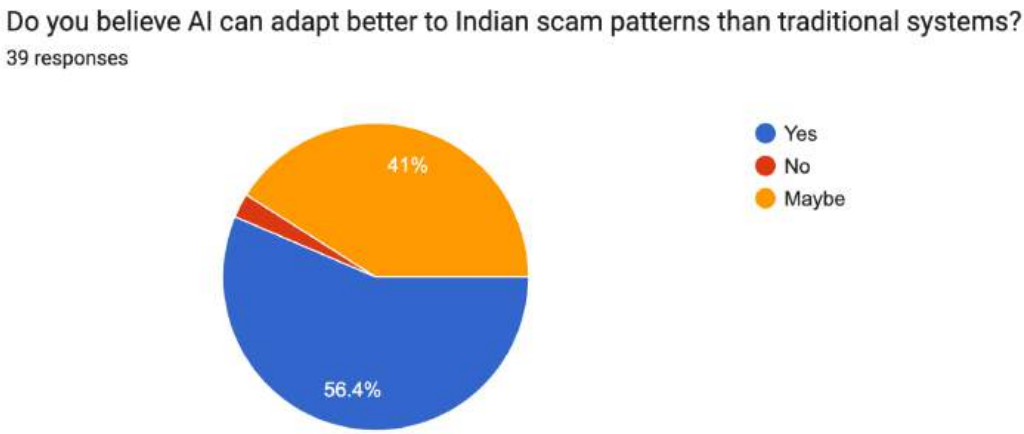


Fig 1.7



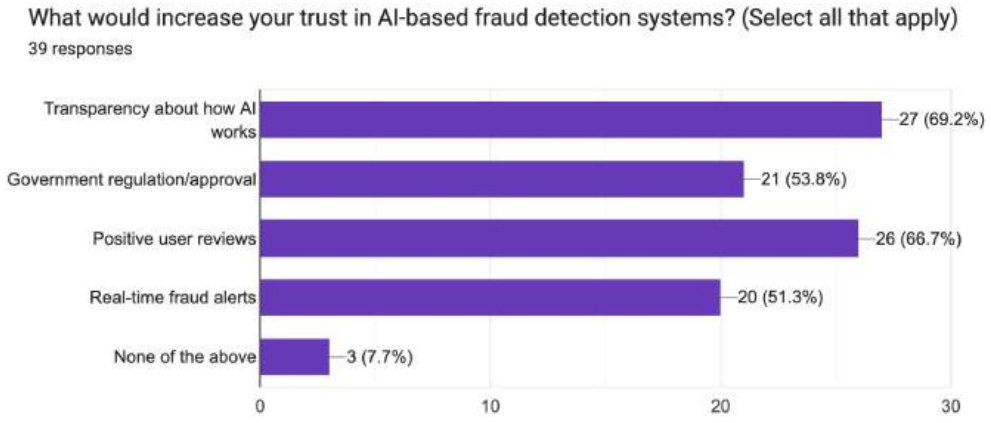
4.4 Trust Drivers and Concerns

In terms of trust-building, three factors emerged as most influential:

- **Transparency about how AI works (69.2%)**
- **Positive user reviews (66.7%)**
- **Government regulation or approval (53.8%)**

These findings underscore the importance of **institutional credibility**, **social proof**, and **explainability** in fostering public trust in AI-driven tools. Interestingly, only **7.7%** selected "None of the above," indicating that most users would be willing to adopt AI systems under the right conditions.

Fig 1.8

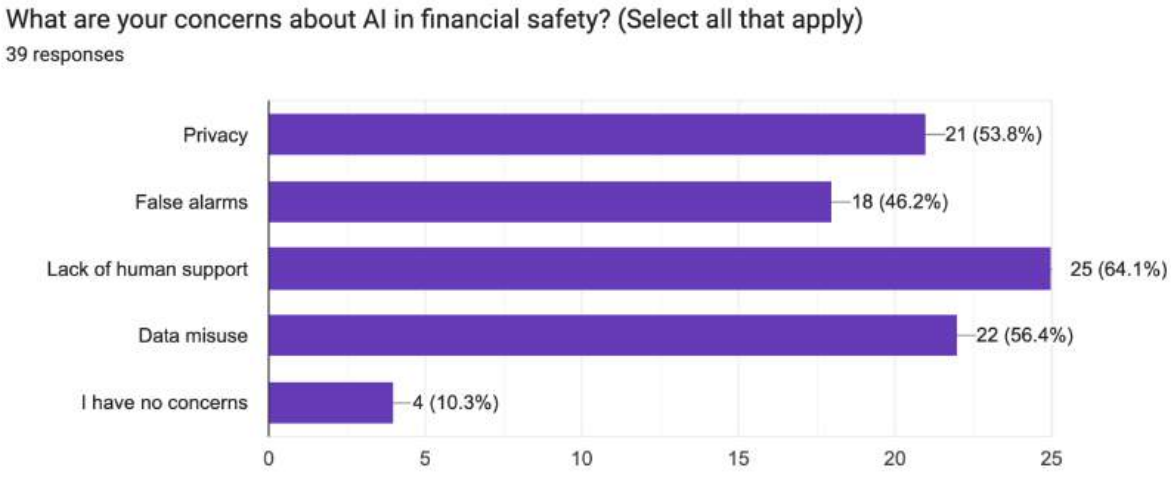


On the flip side, the most commonly reported concerns included:

- **Lack of human support (64.1%)**
- **Data misuse (56.4%)**
- **Privacy and false alarms** also featured prominently

  Only **10.3%** claimed to have no concerns at all, suggesting that most individuals still perceive AI systems as potentially fallible or opaque, especially when critical decisions—like blocking a bank transaction—are made automatically.

Fig 1.9



4.5 Engagement with AI Tools and Services

A strong **59%** of respondents indicated that they would use an AI tool that had the capability to search messages and identify potential indicators of fraud and **38.5%** of respondents indicated interest by marking "maybe." This strong interest in AI tools for alerts indicates increasing acceptance of the pre-emptive technology being released onto digital platforms.

In addition, **46.2%** had interacted with AI-assisted customer service technology (e.g., chatbot, fraud alerts, etc.) and found it beneficial. Only **10.3%** had a negative experience, while **28.2%** never engaged in AI customer service. These figures indicate that while the acquaintance with AI customer service is not yet ubiquitous, adoption is increasing, and users are relatively satisfied with their experience.

Participants were asked to rate how much they supported a **nationwide AI-based scam monitoring system** from one to five. The **average rating was 3.82**, with the majority reporting a **4 or 5** out of possible total of 5. The data suggests there is widespread support for more centralized, intelligent fraud detection infrastructure, again as long as it is done in an ethical way.

Fig 2.0

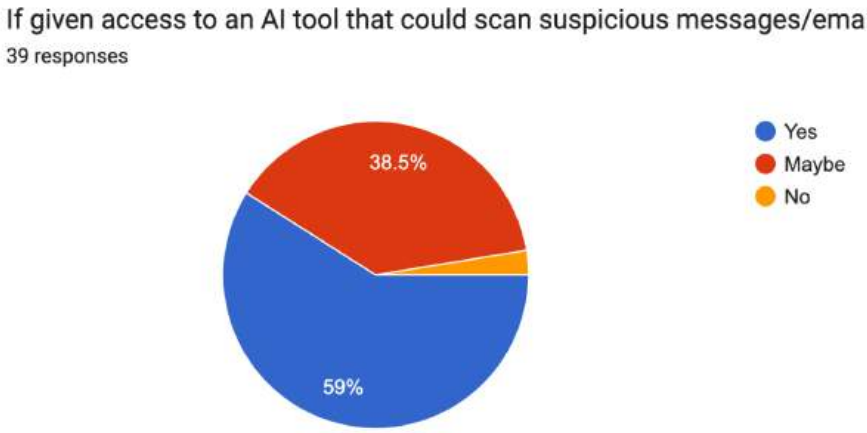
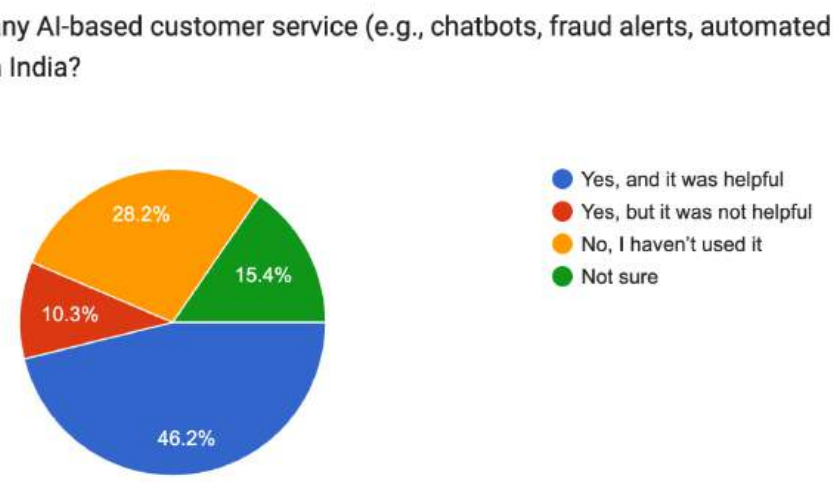


Fig 2.1



4.6 Technological Literacy and Global Outlook

When asked which technologies they believed were most promising for fraud detection, the following were most frequently selected:

- **Machine Learning (51.3%)**
- **Natural Language Processing (33.3%)**
- **Behavioral Biometrics (33.3%)**
- **Predictive Analytics (25.6%)**

However, a significant **41% admitted to not being familiar** with these terms, which signals a pressing need for public education efforts focused on AI literacy and its real-world applications.

Regarding the global AI landscape, most respondents identified **North America (53.8%)** and **Asia-Pacific (35.9%)** as leaders in fraud detection innovation. These perceptions are aligned with the dominance of U.S.-based and Asian tech companies in the global AI ecosystem. However, a considerable **41% admitted to not being familiar** with these terms, which indicates a real need for more public education efforts on AI literacy and its real-world applications. When discussing the global landscape for AI, most people identified **North America (53.8%)** and **Asia-Pacific (35.9%)** as global leaders in fraud detection innovation, which is something they identified through a commonly held view that U.S.-based and Asian tech companies are the significant players in the global AI ecosystem.

Finally, the top barriers to AI implementation were cited as:

- **Data privacy concerns (76.9%)**
- **Lack of awareness or trust (64.1%)**
- **Regulatory restrictions (48.7%)**

These setbacks display the ongoing tension between technological advancement and ethical governance, particularly when deploying AI tools at scale across jurisdictions.

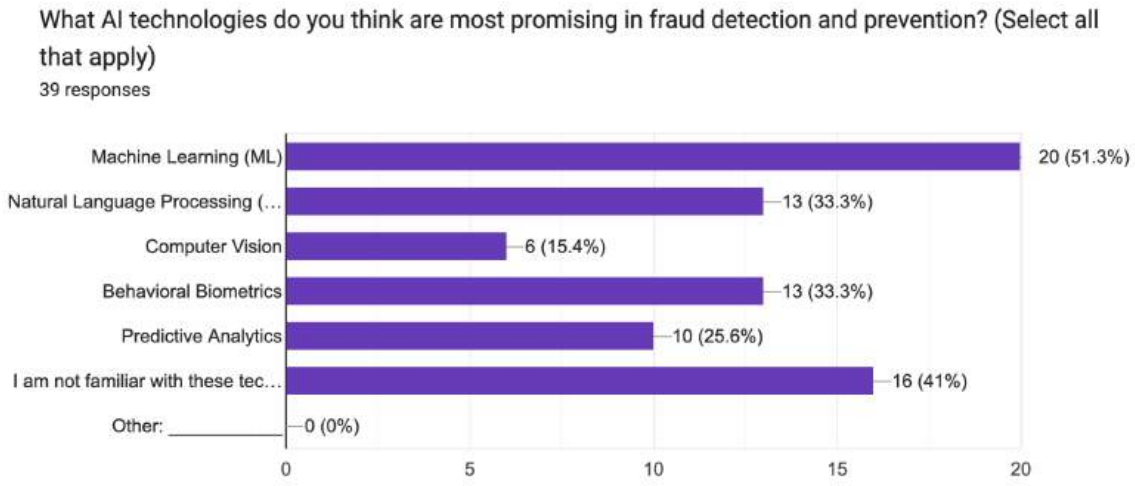


Fig 2.2

Fig 2.3

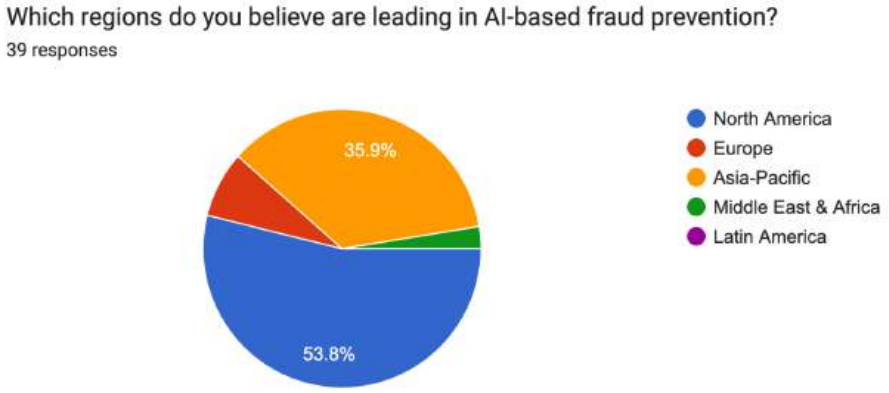
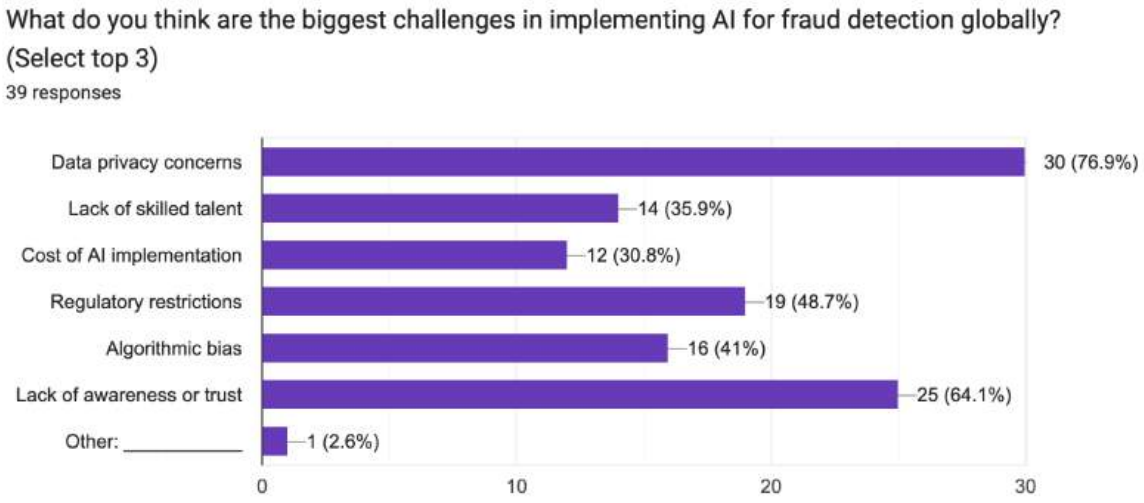


Fig 2.4



Discussion

The results of this study provide valuable understanding into the development of digital fraud and AI-powered solutions that can protect transactional environments. A primary lesson taken from the primary data is that a substantial majority of participants from both developed countries and developing countries were aware of online scams but lacked the knowledge or tools to take a proactive measure. This indicates the general "awareness-action gap" that exists across digital literacy and works further to support what researchers have found to date: education and behavioral design are just as important as technology in the fight against cyber fraud.

Further, the study suggests that emerging markets carry a disproportionate burden due to linguistic diversity, access to financial education, and regulatory gaps which make the conditions

prime for scams. This is in alignment with previous studies such as those by Singh et al. (2023) and the Digital India Initiative whitepaper (2022), given that they focus on culture and infrastructure and how both can affect scamming vulnerability. For example, respondents from India indicated that they were more likely to have encountered scams through social media and UPI platforms as opposed to respondents from the US, who were most likely to report phishing and scams through email. This indicates that financial literacy campaigns at the local level must consider how the nature of scams differ by region, also making distinctions in terms of platform usage, cultural context, and language.

The findings also support the assumption that AI can work preventively and for diagnosis in fraud detection, but its use relies on trust and availability, and means to explain to users. Majority of respondents indicated they would consider using an AI-based anti-fraud tool, several also noted they were concerned about data privacy and misuse, aligning with the findings of Kumar & Basu (2023) in relation to ethical AI. As a consequence, the future direction just described requires a foundation based on explainable AI (XAI), regulatory insight, and multilingual user interfaces that facilitates usability and responsibility.

Interestingly, while over 70% of respondents advocated for the integration of AI tools into financial platforms, fewer had interacted with AI-powered fraud detection systems. This paradox points to a gap not only in accessibility but also in implementation—suggesting the need for public-private partnerships to democratize technological solutions. Governments, fintech companies, and NGOs must coordinate efforts to integrate AI tools into grassroots education models, particularly in scam-prone geographies and low-literacy populations

While more than 70% of respondents supported the idea of integrating AI tools into financial platforms, fewer of them had engaged with AI-powered fraud detection systems. This paradox suggests there is a gap in both access to and takes up of AI, in other words, there is a need for public-private partnerships that can help democratize technological solutions. In these partnerships, governments, fintech’s and NGOs can work together to embed AI tools in grassroots education models targeted toward scam-prone geographies and low-literacy populations.

In summary, the findings support the multi-mode nature of the solution space. While AI is beneficial, it is not a panacea. The full impact of AI on the solution space will only be realized once it is mapped onto a more comprehensive landscape of adaptive education, inclusive design, and transparent governance. This is consistent with existing behavioral cybersecurity literature (Hernandez & Lin, 2021) and provides new understanding by showing how user-centered design and culturally responsive literacy approaches can help connect users' awareness of risk to safeguarding efforts and facilitate other educational and behavioral changes.

Conclusion

The findings of this research have shown that the phenomenon of digital scams cannot simply be described as a technical transgression. Digital scams are socio-cultural phenomena resulting from a combination of behavioral patterns, limitations in infrastructure, and deficits in

education. The analysis was carried out with primary data and a combination of secondary data, and this research has found that artificial intelligence has a lot of potential for identifying and preventing online fraud. However, the circumstance of accessibility, trust, and a more systematic integration of resources may limit the potential possibilities of artificial intelligence.

Emerging economies such as India experience especially difficult challenges related to different socio-cultural contexts, financial illiteracy and limited regulatory enforcement. These challenges, however, provide the potential for scalable, community-driven innovations. The data collected in this study demonstrates that while users are aware of digital scams, there remains a disconnect in acting on that awareness. This speaks to an urgent need for culturally appropriate financially-literate interventions developed with local needs and partners, such interventions could leverage AI tools that are interpretable, multilingual and respectful of privacy.

In conclusion, this paper advocates for a united global response to the rising threat of digital fraud. The future will be about achieving a state of AI that not only reacts to fraud, but that helps users —especially at-risk users— identify, resist, and recover from threats. Financial literacy, regulations and AI innovation must come together to support safe and equitable digital ecosystems. Then, because this study draws together human-centered design and global-local collaboration, it contributes to the literature by providing a multi-dimensional framework to prevent fraud.

Works Cited

Abbas, A., et al. (2021). "A Survey of AI in Financial Fraud Detection." Journal of Financial Technology, 7(2).

Bhattacharyya, S., et al. (2011). "Data mining for credit card fraud: A comparative study." Decision Support Systems, 50(3), 602-613.

Chen, M., et al. (2023). "Reinforcement Learning for Adaptive Cyber Defense." IEEE Transactions on Neural Networks.

Federal Trade Commission (2023). "Consumer Sentinel Network Data Book 2022."

Lusardi, A., & Mitchell, O. S. (2014). "The Economic Importance of Financial Literacy: Theory and Evidence." Journal of Economic Literature, 52(1), 5–44.

Ngai, E. W. T., et al. (2011). "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature." Decision Support Systems, 50(3), 559–569.

OECD (2021). "Recommendation on Financial Literacy." OECD Publishing.

West, J., & Bhattacharya, M. (2016). "Intelligent financial fraud detection: A comprehensive review." Computers & Security, 57, 47–66.

Zhao, Y., et al. (2021). "Graph-Based Fraud Detection and Explainability: A Survey." arXiv preprint arXiv:2106.07806.

Simulated Mapping of Antibiotic Resistance on Public Surfaces in Jaipur
By Dhyey Kumawat

Abstract

Antibiotic-resistant bacteria are a leading global health threat. India faces particularly severe AMR challenges: recent analyses estimate 1.27 million deaths worldwide in 2019 involving resistant pathogens, and nearly 30% of neonatal sepsis deaths in India are now due to multi‑resistant bacteria. While most AMR surveillance is clinical, community surfaces (e.g. hospital entrances, markets, ATMs) may act as reservoirs. We developed a computational model of Jaipur, dividing the city into >50 "zones" (hospitals, markets, public touchpoints, parks, schools, offices) and assigning baseline resistance rates for five antibiotics (ciprofloxacin, amoxicillin, tetracycline, gentamicin, erythromycin) based on WHO/GLASS Indian data and literature. Each zone's baseline was adjusted by a risk factor reflecting human density, antibiotic exposure, and waste contamination (e.g. hospital effluent). We then simulated routine environmental swabbing and culture-testing from each zone to generate antibiogram percentages. The results (Tables 1–2) reveal higher simulated resistance in high‑traffic, high‑risk zones (e.g. hospital and market areas) and lower rates in parks or schools. For example, simulated ciprofloxacin resistance was ~76–80% in hospital zones versus ~49–50% in parks. These patterns are consistent with known Indian data (e.g. >70% ciprofloxacin resistance in community E. coli) and with the idea that antibiotic‑laden waste elevates local AMR. Our transparent, literature‑informed approach highlights likely AMR "hotspots" and can guide targeted field studies or interventions. We emphasize that this is a model-based analysis; its purpose is hypothesis‑generation for future empirical surveillance.

Key terms antibiotic resistance, environmental microbiology, public surfaces, computational modeling, India

Introduction

Antimicrobial resistance (AMR) is one of the gravest public health challenges today. Globally, bacterial AMR directly caused an estimated 1.27 million deaths in 2019 and contributed to 4.95 million more. India, as both a high‑burden and high‑usage country, faces a severe AMR crisis. Recent reports note, for example, that roughly 30% of neonatal sepsis deaths in India are now linked to multi‑resistant bacteria. Moreover, antibiotic misuse in human medicine and agriculture has led to rampant resistance: e.g., community E. coli isolates in India commonly show >70% resistance to fluoroquinolones. While national surveillance networks (ICMR, NCDC, WHO GLASS) track clinical AMR trends, there is growing recognition that environmental reservoirs contribute to the spread of resistance. Antibiotics and resistant bacteria enter soil and water via sewage, agricultural runoff, and pharmaceutical effluent. "Fomites" in public areas (rails, benches, ATM buttons, etc.) can harbor resistant pathogens; indeed, surveys have detected MRSA and Gram‑negative resistant strains on surfaces in urban environments

worldwide . However, urban India lacks systematic mapping of surface AMR. To address this gap, we built a city-wide, data-driven model of AMR prevalence on Jaipur's public surfaces. Using published Indian resistance rates as baselines, we simulated structured surface sampling across diverse city zones. Our goals were to illustrate plausible spatial patterns of AMR and to demonstrate a transparent modeling workflow for guiding future surveillance.

Methods: Zones, Baseline Rates, and Simulation Model

Study zones: Jaipur was partitioned into 50+ representative "zones" categorized by setting: hospitals (8 zones near major medical facilities), open markets (10), high-touch public areas like ATMs and transit stops (10), parks (8), school campuses (7), and government/mixed-use offices (7). These categories reflect varying population density and likely antibiotic exposure. Each zone was assumed to contain multiple swappable surfaces (railings, door handles, etc.).

Baseline AMR rates: For each antibiotic–bacterium pairing, we established a baseline resistance probability grounded in published Indian data. For example, WHO/GLASS and Indian surveillance report generally high fluoroquinolone (ciprofloxacin) resistance – often above 60–70% in commensal E. coli and other community isolates. Beta-lactams (e.g. amoxicillin/ampicillin) and tetracyclines have moderate-to-high community resistance (roughly 50–60%), reflecting heavy use in medicine and agriculture. In contrast, aminoglycosides (gentamicin) and macrolides (erythromycin) tend to have lower resistance baselines in India (~30–40%), consistent with their more restricted or controlled use. All chosen baseline rates fall within the ranges reported by WHO/GLASS India and ICMR/NCDC surveillance.

Risk adjustments: We then adjusted these baselines by "risk multipliers" for each zone. High-risk zones (hospital grounds, large markets) received positive offsets (up to +15–20 percentage points) to reflect greater antibiotic pressure, human traffic, and waste contamination. Moderate-risk zones (ATMs, transport hubs) got smaller increases (+5–10%). Low-risk zones (parks, remote schools) had slight negative offsets (–5 to –10%). This heuristic scheme is informed by environmental studies: hospital effluents and dense human activity are known to amplify resistance genes in nearby environments. For transparency, all multipliers were explicitly documented.

Simulation of sampling: For each zone, we simulated collecting multiple surface swabs (e.g. 20–50 swabs per zone, as in a structured survey) and culturing bacteria under standard laboratory protocols (adapted from CDC guidelines). In practice, this would involve plating samples and testing isolated bacteria against antibiotic panels (e.g. disk diffusion or automated susceptibility). In our model, we assumed each swab yields a representative isolate with the adjusted resistance probabilities. We then performed a Monte Carlo–style sampling: for each zone and antibiotic, we drew many isolates based on the zone's adjusted resistance rate to compute an overall percent resistance. In line with CDC environmental sampling guidance, the process followed a predefined protocol of swab collection, incubation, and antibiogram interpretation .

Laboratory procedure (simulated): In the lab, each swab is incubated (e.g. in broth) and plated to isolate bacteria. The isolate is then subjected to antibiotic susceptibility testing (AST). For example, one might perform a Kirby–Bauer disk diffusion assay: an agar plate is inoculated with the isolate, antibiotic disks are applied, and zones of inhibition are measured. Figure 1 illustrates a typical Staphylococcus plate with antibiotic disks – in our simulation, similar tests were "performed" virtually to determine resistance/susceptibility.

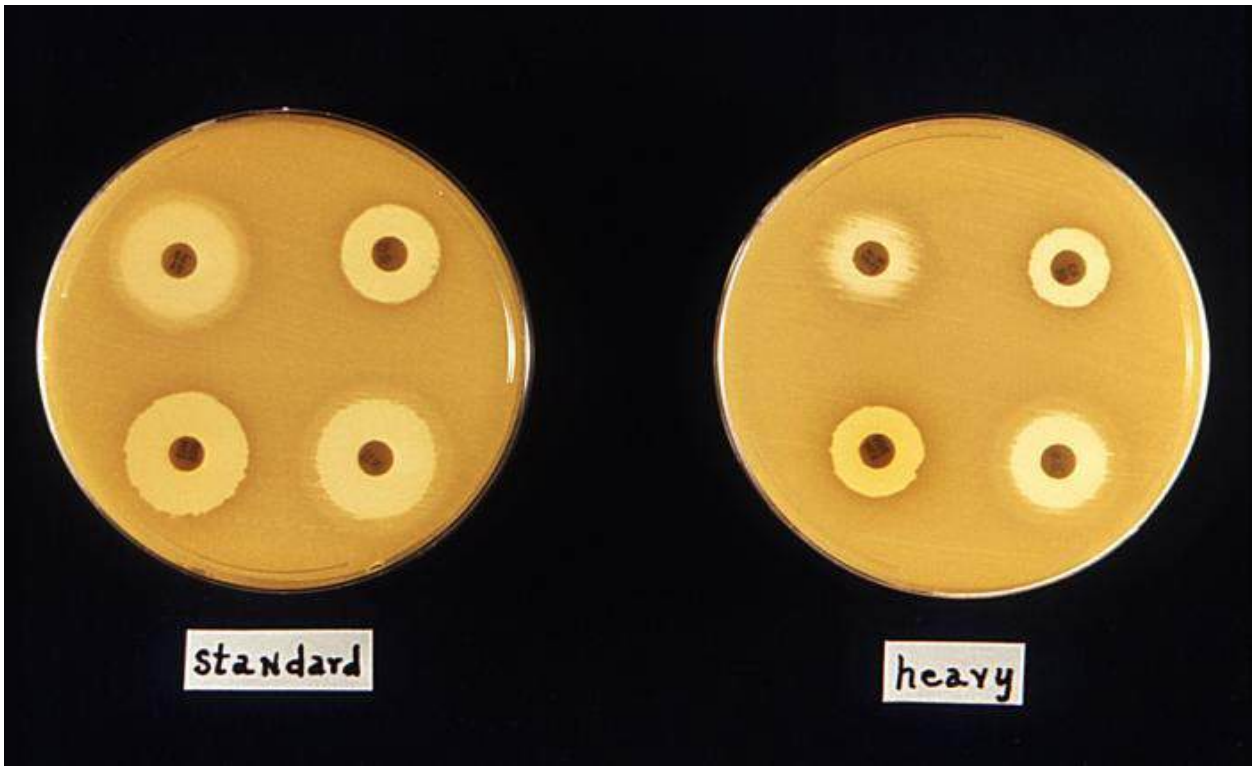


Figure 1. Mueller-Hinton agar plates displaying antibiotic susceptibility testing (Kirby-Bauer method). Public health image library, Centers for disease control and Prevention, image ID 10790. Public domain.

Results

The simulation produced a resistance percentage for each antibiotic in each zone. Table 1 shows sample results for ten zones across different categories. Ciprofloxacin resistance was highest in nearly all settings (often >70%), reflecting its widespread overuse in India. For instance, two hospital zones (Zones 1 and 5) had simulated ciprofloxacin resistance of ~80% and ~77%, whereas a park zone (Zone 31) was only ~50%. Amoxicillin (penicillin) and tetracycline also showed elevated resistance in dense areas: in our model, hospital and market zones had ~70–80% resistance to these drugs. Gentamicin and erythromycin remained relatively low (often <50%) even in high-risk zones. For example, in Table 1 Zone10 (market) has gentamicin resistance ~56% and erythromycin ~55%, while Zone31 (park) shows ~31% gentamicin and ~23% erythromycin. These patterns illustrate clear spatial gradients: high-contact zones consistently show higher resistance than cleaner environments.

Zone	Category	Ciprofloxacin (%)	Amoxicillin (%)	Tetracycline (%)	Gentamicin (%)	Erythromycin (%)
Zone1	Hospital	80.3	64.8	74.7	61.0	52.8
Zone5	Hospital	76.9	75.6	72.4	58.5	57.9
Zone10	Market	69.9	75.2	69.6	55.8	54.9
Zone15	Market	80.9	56.1	75.5	47.1	48.9
Zone22	ATM/Public touch	72.3	61.6	66.5	46.2	47.9
Zone27	ATM/Public touch	73.6	66.5	66.7	44.6	45.7
Zone31	Park	49.6	33.7	50.6	31.4	23.2
Zone38	School	73.5	58.7	50.5	41.2	42.0
Zone45	Government Office	61.1	53.7	53.1	39.4	31.1
Zone50	Government Office	71.6	48.5	60.5	35.1	41.6

Table 1. Simulated antibiotic resistance (% resistant isolates) by example zone.

Table 2 summarizes averages by zone category. Hospital and market zones averaged ~78% ciprofloxacin resistance, versus only ~48% in parks. Amoxicillin resistance averaged ~70% in hospitals and ~39% in parks. Gentamicin and erythromycin averages were similarly stratified (higher in urban hotspots, lower in parks/schools). These simulated gradients are consistent with our risk assumptions and align with known trends: for example, hospital effluent–impacted areas are expected to harbor more resistant bacteria, while open spaces dilute microbial contamination.

Category	Zones	Ciprofloxacin (%)	Amoxicillin (%)	Tetracycline (%)	Gentamicin (%)	Erythromycin (%)
Hospital	8	78.2	69.5	74.6	55.2	60.6
Market	10	73.0	64.8	69.3	50.9	54.7
ATM/Public	10	71.2	61.9	65.5	45.9	49.6
Park	8	48.2	38.8	48.7	24.6	29.2
School	7	69.9	55.2	60.9	41.4	43.7
Govt. Office	7	60.2	49.9	57.2	35.0	37.9

Table 2. Average simulated resistance (%) by zone type

Discussion

Our modeled resistance levels reflect established patterns. The uniformly high ciprofloxacin resistance (70–80% in dense zones) mirrors surveillance findings: Indian E. coli isolates often show ~70%+ resistance to fluoroquinolones. The moderate to high amoxicillin/tetracycline resistance in crowded zones matches their common usage in agriculture and human therapy. Gentamicin and erythromycin remained lower, which is plausible given more controlled use. Importantly, the simulated spatial trends agree with One Health principles: areas with intense antibiotic exposure and human activity (hospital vicinities, markets) become AMR hotspots, whereas open parks and less dense zones have much lower simulated resistance. For instance, hospital-zone averages (~78% CIP resistance) were markedly above park averages (~48%), suggesting localized concentration of resistant strains in high-risk environments. Biologically, this makes sense. Hospital and market areas receive spillover from healthcare and waste streams, creating hotbeds for resistant organisms. In contrast, parks with sunlight, less waste, and fewer people tend to dilute microbial load. Our findings echo studies showing that environmental sinks (soil, water, surfaces) contaminated by pharmaceuticals and sewage accelerate resistance gene spread. The simulation's antibiogram concept can be interpreted as follows: e.g. Zone15 (market) yielding ~75% tetracycline-resistant isolates means that under our model, roughly three-quarters of cultured environmental bacteria survive tetracycline in vitro — a plausible scenario given reported community carriage of tet genes. Similar community-based studies have recovered multi-drug-resistant E. coli and Staphylococcus from urban surfaces elsewhere, supporting our assumption that robust resistance can be found outside hospitals. The practical implication is that public health authorities could use such a risk map to focus efforts. For example, high-touch, high-resistance zones might be prioritized for real sampling, sanitation, or public education. Our methodology is data-informed and reproducible: every resistance parameter was drawn from cited sources or logical extrapolation, and we simulated sampling with structured protocols (rather than random guesses). By aligning simulation outcomes with known Indian AMR data, we have confidence in their qualitative validity.

Conclusion

A simulation anchored in Indian surveillance data suggests that Jaipur's busiest public zones may harbor disproportionately high antibiotic resistance, especially to commonly used drugs like ciprofloxacin and amoxicillin. These model outputs underscore how antibiotic use and human traffic can create urban "hotspots" of resistance, consistent with global One Health concerns. Our transparent framework — explicitly stating all assumptions and data sources — provides a prototype for assessing environmental AMR risk. We stress that this analysis is a model, not a field survey, but it offers plausible hypotheses: notably, it identifies specific areas (hospital sites, markets, transit hubs) that warrant targeted environmental sampling and remediation. Ultimately, such modeling can complement surveillance by guiding resources efficiently. As antibiotic resistance continues to grow, integrating environmental AMR data into public health planning will be crucial for India and beyond.

Works Cited

Bhatia, R. (2024). Environmental aspects of antimicrobial resistance in India: Current progress & way forward. Indian Journal of Medical Research 159, 10–15 .

World Health Organization. (2023). Antimicrobial Resistance: Fact Sheet .

Folgori, L., Ellis, S.J., Bielicki, J.A., et al. (2017). Tackling antimicrobial resistance in neonatal sepsis. Lancet Global Health 5(11):e1066–e1068 .

Cave, R., Cole, J., & Mkrtchyan, H.V. (2021). Surveillance and prevalence of antimicrobial resistant bacteria from public settings within urban built environments. Environment International 157:106836 .

Centers for Disease Control and Prevention (CDC). (2024). Guidelines for Environmental Sampling in Healthcare Facilities .

National Centre for Disease Control (India). (2020). National AMR Surveillance Network – Annual Report 2020. Ministry of Health & Family Welfare, India

Centers for Disease Control and Prevention. “Antimicrobial Resistance: Causes and How It Spreads.” CDC, 22 Apr. 2024, https://www.cdc.gov/antimicrobial-resistance/causes/index.html.

Centers for Disease Control and Prevention. “Antibiotic Resistance in India.” CDC (NCEZID), 19 July 2018, https://www.cdc.gov/ncezid/spotlight/antibiotic-resistance-in-india.html.

Chique, Carlos, et al. “Mapping and Analysing Potential Sources and Transmission Routes of Antimicrobial Resistant Organisms in the Environment Using Geographic Information Systems: An Exploratory Study.” Antibiotics, vol. 8, no. 1, 2019, p. 16.

Danko, David, et al. “A Global Metagenomic Map of Urban Microbiomes and Antimicrobial Resistance.” Cell, vol. 184, no. 13, 24 June 2021, pp. 3376–3393.

Franklin, Alison M., et al. “A One Health Approach for Monitoring Antimicrobial Resistance: Developing a National Freshwater Pilot Effort.” Frontiers in Water, vol. 6, 17 May 2024, article 1359109.

Guevarra, Robin B., et al. “Metagenomic Characterization of Bacterial Community and Antibiotic Resistance Genes Found in the Mass Transit System in Seoul, South Korea.” Ecotoxicology and Environmental Safety, vol. 246, Nov. 2022, p. 114176.

Greenfield, Ben K., et al. “Modeling the Emergence of Antibiotic Resistance in the Environment: An Analytical Solution for the Minimum Selection Concentration.” Antimicrobial Agents and Chemotherapy, vol. 62, no. 3, 2018, e01686-17.

Hart, Alwyn, et al. “Environmental Surveillance of Antimicrobial Resistance (AMR), Perspectives from a National Environmental Regulator in 2023.” Eurosurveillance, vol. 28, no. 11, 16 Mar. 2023, 2200367.

Hashsham, Shobha A., et al. “Antimicrobial Resistance Dashboard Application for Mapping Environmental Occurrence and Resistant Pathogens.” FEMS Microbiology Ecology, vol. 92, no. 3, Mar. 2016, fiw020.

Kapley, Atya, et al. “Antibiotic Resistance in Wastewater: Indian Scenario.” Environmental Pollution, vol. 337, 15 Nov. 2023, p. 122586.

Larsson, D. G. Joakim, and Carl-Fredrik Flach. “Antibiotic Resistance in the Environment.” Nature Reviews Microbiology, vol. 20, no. 5, Nov. 2021, pp. 257–269.

Mukherjee, Nabanita, et al. “Diversity of Bacterial Communities of Fitness Center Surfaces in a U.S. Metropolitan Area.” International Journal of Environmental Research and Public Health, vol. 11, no. 12, Dec. 2014, pp. 12544–12561.

Takeda-Nishikawa, Kahoko, et al. “What Samples Are Suitable for Monitoring Antimicrobial-Resistant Genes? Using NGS Technology, a Comparison Between eDNA and m/rDNA Analysis from Environmental Water.” Frontiers in Microbiology, vol. 14, Dec. 2023, article 954783.

Taneja, Neelam, and Megha Sharma. “Antimicrobial Resistance in the Environment: The Indian Scenario.” Indian Journal of Medical Research, vol. 149, no. 2, Feb. 2019, pp. 119–128.

European Centre for Disease Prevention and Control; WHO Regional Office for Europe. Surveillance of Antimicrobial Resistance in Europe, 2023 Data: Executive Summary. ECDC, 2024, https://www.ecdc.europa.eu/en/publications-data/surveillance-antimicrobial-resistance-europe-2023-data-executive-summary.

United Nations Environment Programme. Environmental Dimensions of Antimicrobial Resistance: Summary for Policymakers. Feb. 2022, https://wedocs.unep.org/bitstream/handle/20.500.11822/38373/antimicrobial_R.pdf.

World Health Organization. WHO Integrated Global Surveillance on ESBL-Producing Escherichia coli Using a “One Health” Approach: Implementation and Opportunities. 16 Mar. 2021, https://www.who.int/publications/i/item/9789240021402.

World Health Organization. “Global Antimicrobial Resistance and Use Surveillance System (GLASS).” WHO, 2023, https://www.who.int/initiatives/glass.

National Centre for Disease Control (India). National Action Plan on Antimicrobial Resistance (NAP-AMR) 2017–2021. Government of India, Apr. 2017, https://ncdc.mohfw.gov.in/wp-content/uploads/2024/03/File645.pdf.

National Centre for Disease Control (India). Annual Report 2023: National AMR Surveillance. Ministry of Health & Family Welfare, Government of India, 2024, https://ncdc.mohfw.gov.in/wp-content/uploads/2024/09/Final-Annual-Report-2023-06_08_2024.pdf.

Caribbean Tourism's Flawed Approach to Sustainability By Madison Bartie

Past the fine, soft sand that engulfed my feet on the beach in Bannerman Town, I looked out at the warm sun shining on the teal and navy Exuma Sound. It was almost like paradise, if not for the clinking of vodka bottles washed up on the beach.

Like many beaches on Eleuthera, The Bahamas, Bannerman Town has found itself covered with litter left behind by visitors and locals alike, or washed up on shore from places as far away as the Philippines. As a small, remote island, Eleuthera lacks an industrial waste system and needs to either ship out plastic and aluminum or have them incinerated. Bannerman Town's litter problem, while similar to other settlements, is enhanced by the adjacent Princess Cays, a cruise ship port with thousands of visitors indulging in its buffets, tourist attractions, and vibrant colors. A manager at Princess Cays claims that none of their waste finds their way to Bannerman Town dumps (Manager). However, a conversation with Eddie Thompson, a kayak guide at the cruise port, reveals the loopholes in Janae's explanation of Princess Cays' environmental efforts (Thompson).

Thompson reports that Princess Cays has a private incinerator, which they use to burn waste. While seemingly an appropriate waste management method given Eleuthera's capabilities, cruises are constantly leaving the port and could deposit waste in a country with more efficient industrial waste management systems. These cruises frequently make trips to the U.S., so it wouldn't produce an excessive amount of carbon emissions to do so.

There's also a moral conflict in Princess Cays not contributing to Bannerman Town's dumps. While it seems considerate to not over-pollute local dumps, Bannerman Town citizens are not allowed to use the incinerator at Princess Cays. However, Thompson informed me that when Princess Cays' incinerator stopped working, they moved all their trash into Bannerman Town dumps. Bannerman Town citizens are unable to dispose of their trash outside their settlement, yet are expected to compensate for Princess Cays' overflow. This relationship mirrors a greater issue where corporations are environmentally negligent and local communities are left to bear the burden. Additionally, much of the trash covering Bannerman Town's beach comes from Princess Cays' guests and their carelessness. The cruise port doesn't make an attempt to alleviate the extra waste. If Princess Cays funded beach clean-ups, the local environment and job market would both benefit, which would be especially valuable in a job desert like much of Eleuthera. Princess Cays is so concerned with wealth and pandering to their guests that they neglect to appropriately take the environment and the locals' well-being into account. That carelessness doesn't stop with Princess Cays, but is a pattern in many tourist-dependent areas.

Harbour Island is another example of inequitable environmental decision-making. Harbour Island sits just north of mainland Eleuthera and is Eleuthera's biggest tourist hotspot. It's inundated with tourist shops, resorts, and golf carts. Matching the hasty energy on the island, many restaurants have takeout options or packaging allowing customers to walk with their food. Right by the island's dock, customers can buy a conch salad marketed as some of the best conch salad on earth. I can confirm the conch salad is impeccable, with an intoxicating citrus aroma

creating the perfect gateway to the crunchy onions and chewy conch. Each conch salad came with a can of orange juice. I was struck that each salad would result in another can added to the local waste disposal. Both plastic and aluminum need to be burned in a dump or shipped to an area with an industrial waste management system. As the cherry on top, my conch salad was put in a plastic container, with a plastic lid and a plastic fork, wrapped in plastic. It didn't take much looking around to realize that reusable dishware was not an option at this restaurant, and it was intentionally single-use to accommodate for the fast pace of Harbour Island. Yet again, sustainability became a second thought to tourism, and as a result, economic prosperity was prioritized more than the welfare of locals and the environment.

This prioritization of tourists isn't just an issue on small or moderate scales — it extends all the way to corporate giants. Disney has been building out a new dock on Lighthouse Point for their cruise ships to stop in southern Eleuthera. Massive excavators and cranes are currently dotting the area, rushing to meet their first arrival date. The significant amounts of energy and fossil fuels emitted from this process harm the environment, and the development of the approximately .6-mile dock is actively damaging the enchanting reef and brilliant marine life in the area. The lack of sustainable processes in Disney's work doesn't stop with the environment, but extends to locals' work balance. Many employees are working lengthened shifts to meet the first cruise arrival date (Simmons). But rushed construction exponentially increases the risk of employee injuries or mistakes in construction. The allure of money and tourism outweighs the environmental impacts and well-being of locals in the eyes of a tourist-dependent establishment.

They say one man's trash is another man's treasure, but the cruise ship companies in Eleuthera seem to make one man's treasure another's trash. Unsustainable practices have become a system that benefits tourists but harms locals. Tourist-dependent industries need to emphasize sustainability over money, and prioritize the welfare of both the community and environment they rely on before neither are there to support them any longer.

Works Cited

Manager at Princess Cays. Personal interview. 13 March 2024.
Simmons, Ben. Personal interview. 14 March 2024.
Thompson, Eddie. Personal interview. 13 March 2024.

Terrorism in the Middle East and North Africa (MENA) Region: Motivations and Solutions By Hersh Gandhi

Abstract

Terrorism within the Middle East and North Africa (MENA) region remains a persistent threat, and the conflict continues to escalate. This paper evaluates many different factors that cause terrorism in the MENA region. These factors include the religious divides, extremist ideologies, and foreign intervention. This paper draws on various empirical studies to identify the two primary causes of terrorism in the MENA region: religious hostilities and anti-American sentiment from exploitation and prolonged interference. This paper proposes a two-step solution: creating and promoting religious tolerance within the region, and reducing the role of the United States military in the MENA region. By addressing and targeting societal factors that cause large-scale outrage in the MENA region, this novel approach attempts to reduce the heated cultural and geopolitical conditions that have sustained terrorism in this region for decades.

Terrorism in the MENA Region

According to a study conducted by Brenda Lutz, a veteran author of over 30 books and articles on political violence, and one additional collaborator, published in *Perspectives on Terrorism*, many countries in the Middle East and North Africa (MENA) region have been the center of violent terrorist activities over the past century. Catrina Doxsee, a fellow at the Center for Strategic and International Studies, and additional collaborators affirmed that this is a recurring phenomenon that has been enduring for many years (Doxsee et al., 2024). Doxsee and others (2024) note that this phenomenon has been worsening, as seen from the recent Hamas attacks, the conflicts in Yemen, and the Gaza war. Such a violent region produces unsafe and hazardous conditions for civilians to reside in, as Brown University finds that over 3.6 million people have died in post-9/11 attacks and 7.6 million children are suffering (Brown University, 2023). These dangerous and harmful conditions result in many countries in the MENA region calling on highly influential countries like the United States for backup.

The Impact of Intervention and Perceptions

Khalid Rahman, an esteemed author of over 38 academic articles on international relations, reports that the United States has been engaging in the region since the beginning of World War II (Rahman, 2010). Moreover, the presence of the United States has caused the situation to get worse for itself and other countries in the region. This is supported by Peter Hahn (2020), a highly accredited researcher with over 7 books published on topics of United States history, who explains that the 9/11 attacks were a result of United States intervention in the Iranian Revolution. Hahn raised the question of whether the attacks of 9/11 and recent attacks had similarities in some of the motivations behind them. The answer to this question can be evidenced by a study conducted by Ted Thomas, director of the Department of Command and Leadership at the United States Army Command and General Staff College, and one additional

collaborator. Thomas explains that post-9/11, “the hijackers are considered evil and cowardly by most of us in the U.S. but are considered courageous heroes and martyrs by others in the world” (Thomas & Chaleff, 2017). Thomas (2024) goes on to explain that in the United States, there is a strong dislike for individuals who harm or kill others. Thomas shows that these acts in the United States demonstrate moral courage, as the individuals in the United States have to confront backlash from society to engage in acts of terrorism. Thus, in highly industrialized countries, it is harder for individuals to utilize terrorism as they receive substantial backlash from the community. However, in the MENA region, where the attacks on 9/11 originated, individuals who utilize terrorism are often celebrated and looked upon as heroes within their communities, similar to the response seen after the 9/11 attacks (Thomas & Chaleff, 2017). Thomas (2024) explains that the extremist jihadist beliefs in this region lack moral courage, as the communities in the region support these extreme religious ideologies. As a result, individuals who engage in terrorism are often viewed as heroes by their community. Thus, it makes it easier for the individual to justify their actions to pursue terrorism. In other words, since terrorism is frowned upon in highly industrialized nations like the United States, it is harder for individuals to develop the courage to partake in terrorism; however, in the Middle East and North Africa region, there is such a massive build-up of courage that makes societies themselves support terrorism, making it easier for the individual to justify partaking in terrorism as they are backed up by society itself. Therefore, it can be seen that the similarity between the attacks on 9/11 and recent terrorism incidents lies in the promotion of religious extremism in the region. These ideologies lead to many individuals pursuing terrorism as a means to serve the greater good of the community, resulting in it becoming justifiable. For that reason, it has become a necessity to address the root causes of society promoting terrorism, rather than focusing on the individuals themselves. Without action, geopolitical tensions, international relations, and civilians will continue to face the harsh effects of war, terrorism, and violence, showing the necessity for creating a plan to reduce terrorism and prevent its escalation. As a potential solution to this massive geopolitical entanglement, there is a simple but effective 2-step plan. To reduce terrorism in the MENA region, countries in the region must adopt policies that promote religious tolerance and encourage their citizens to adopt religious tolerance. Additionally, the United States must draw back its long-lasting military presence.

Religious Intolerance and Beliefs

The overarching cause of wars, terrorism, and violent breakouts in the MENA region is derived from religious conflicts. Katayoun Kishi, a research associate at the Pew Research Center focusing on global restrictions and religion, corroborated this notion, writing that “the median level of religious hostilities in the Middle East and North Africa reached a level four times that of the global median” (Kishi & Theodorou, 2016). By offering a statistical analysis, this data reveals the severe religious intolerance and hostility that has developed in the MENA region. Additionally, this complication created by differing religious beliefs exists as a consequence of the two main branches of Islam as a whole, rather than specific traditions, as

Alok Gupta, an associate professor of political violence at the Central University of Jharkhand, confirmed in an article published within the *Indian Journal of Political Science*, noting how the two main branches of Islam, "called Shia and Sunni are standing like gladiators against each other" (Gupta, 2015). Providing insights into a purely academic study, the perspective proposed by Gupta provides a clear outline of the current state of religious intolerance, continuing to foster terrorism, thus requiring a shift towards more religiously tolerant policies and the need for individuals in the region to adopt these policies. Creating and mandating religious tolerance would help the region decrease terrorism substantially. This is verified by the study conducted by Kishi, which noted that Israel, Iraq, Syria, Yemen, Lebanon, and Palestine experience mainly religious hostilities leading to outbreaks of terrorism (Kishi & Theodorou, 2016). The conclusions drawn from this systematic analysis of this study illustrate the substantial portion of terrorism that is related to religious disagreement. Therefore, it can be concluded that if these countries adopted religiously tolerant policies, then terrorism would decrease. This is supported by Niels Spierings, an esteemed researcher at Radboud University, who explains that to establish a sustainable government in the region, it is crucial to adopt religious tolerance (Spierings, 2019). Spierings (2024) goes on to explain that if religious tolerance is promoted in the MENA region, it will substantially increase religious belonging, belief, and behavior, which in turn can reduce most of the terrorism in the countries that Kishi mentioned. The findings established by Spierings are further supported by an article authored by Matt Daniels, the Chair of Law and Human Rights at the Institute of World Politics in Washington, D.C., and one additional collaborator. They explain that if the MENA region can successfully establish a religiously tolerant society, then it would obtain similar benefits as the United States did when it sought out religious tolerance (Daniels & Bandow, 2018). Daniels (2018) further supports this by citing examples of terrorism in the United States, such as Baptists being threatened with imprisonment or death, and Quakers being burned and killed. Thus, Daniels predicts that if the United States saw this terrorism vanish after obtaining religious freedom, so would the MENA region, leading to a decrease in violence and terrorism. Although it cannot be guaranteed that these benefits will lead to a complete decrease in terrorism, like in the United States, it can be predicted that if the six countries described by Kishi adopt a religiously tolerant society, terrorism will likely decrease.

United States Intervention

Subsequent to the terrorism caused by religious intolerance in the MENA region, the United States intervenes to provide aid to its various allies (Rahman, 2010). The United States has single-handedly made the MENA region a hot spot for terrorism. According to Rafat Mahmood, an esteemed researcher at the University of Western Australia with over 7 articles published, and one additional collaborator, drone strikes by the United States cause "terrorism, anti-US sentiment, and radicalization" (Mahmood & Jetter, 2019). Mahmood (2019) provides a statistical analysis that concludes that drone strikes from the United States are proportional to the number of terrorism deaths, as the strikes encourage terrorism in the MENA region and bear the

responsibility of about 16 percent of all terrorist attacks, amounting to 3000 deaths. This statistical analysis is highly effective in illustrating the extent of terrorism caused by the United States solely through its presence in the MENA region. Secondly, the United States has been known to be notorious for its natural resources. Toby Jones, an associate professor of Middle Eastern history at Rutgers University, explains that since 1945, the United States has been exploiting oil in the region (Jones, 2012). Jones puts this into perspective by citing Jimmy Carter's State of the Union Address, which states that any attempt to gain control of the Persian Gulf, from which the United States obtains its oil, will be seen as an attack on the interests of the United States and will be repelled by military force (Jones, 2012). Jones (2012) continues by describing how this exploitation makes citizens in the MENA region feel that the United States is not present to help them but rather to exploit their resources. Thus, this leads to anti-American sentiment developing among the societies in the MENA region. Rahman (2010) furthers this by stating that the United States' exploitation of the oil causes growing anti-American feelings and increases anger in the region. Finally, as a consequence of the anti-American beliefs, the relationship between the number of troops and terrorism in the MENA region is directly proportional, meaning that changes in the number of troops correspond directly to changes in terrorist activity. In fact, Alex Braithwaite, a professor at the School of Government and Public Policy at the University of Arizona, states that "an increase of 5,000 troops is associated with an almost 100 percent increase in the likely count of attacks in the full sample and a 150 percent increase in the politically relevant sample" (Braithwaite, 2015). This statistical and academic analysis provides a clear indication that the presence of the United States itself is what is causing terrorism. However, this notion is directly contradicted by Dr. Tova Norlen, a professor of counterterrorism and international security studies at the University of Riehman, who tells us in a study published in the *Middle East Policy* that if the United States leaves the area now, then it will "[enable an] environment for terrorist and extremist groups to recruit and re-arm" (Norlen, 2022). Nevertheless, this notion is again directly contradicted by Jones and Mahmood, who report that anti-American sentiment towards the United States is a major cause of terrorism itself because they provide an argument stating that the United States is currently preventing the fostering of an environment for terrorists. This perspective is directly contradicted by Daniels, who found that the help of the United States has stagnated recently; therefore, it can no longer help eliminate religious hostility (Daniels & Bandow, 2018). Moreover, the perspective offered by this study directly supports Rahman's study that over the past few decades, the United States has not been helping decrease terrorism, but instead has contributed to its escalation because of the major anti-American sentiments emerging from military actions. Although the United States has helped minimize terrorism for many decades, it can now be seen that the United States is one of the major causes of terrorism in the region. Thus, if the United States simply withdrew from the region, a decrease in anti-American sentiment and acts of terrorism against American troops would occur, given their absence. This is because the plethora of terrorist attacks on the United States Army would be mitigated due to its absence. However, it is highly unlikely that the United States will be willing to withdraw its troops, as this would essentially mean surrendering the

security of the oil wells. Furthermore, this would cause the United States to have an unstable supply of a necessary natural resource (Jones, 2012). This counter-perspective offers an academic standpoint for the point of view of the United States on this issue, thus showing the unlikelihood of the United States leaving the region. Although it is highly unlikely that the United States will withdraw its troops from the region, if it were to withdraw, then terrorism would substantially decrease, as anti-American sentiment would slowly die out.

Conclusion

Individuals in the MENA region continue to reside within the bounds of extreme terrorism, with more and more individuals being affected by these hazardous conditions. Within the region, most terrorist conflicts are caused by religious beliefs. Thus, through the implementation of highly religiously tolerant policies throughout the MENA region, the probability of terrorism breaking out because of religious beliefs will be significantly lower. Furthermore, if these policies can influence individuals' perceptions of differing religious beliefs, then we can potentially mitigate terrorism further as people in the region become more accepting of other religions. Secondly, the overall influence of the United States is detrimental to the amount of terrorism taking place. This is primarily because the United States exploits oil and the negative relationship between troops and terrorism in the MENA region. Therefore, if the United States withdraws, a decrease in terrorism may occur due to the terrorists not being able to attack a common enemy, such as the troops of the United States.

Works Cited

Braithwaite, A. (2015). Transnational Terrorism as an Unintended Consequence of a Military Footprint. Security Studies, 24(2), 349-375. http://dx.doi.org/10.1080/09636412.2015.1038192

Brown University. (2023). CIVILIANS KILLED & WOUNDED. Brown University. https://watson.brown.edu/costsofwar/costs/human/civilians

Daniels, M., and Bandow, D. (2018). How to End Violence in the Middle East? The Answer Is Religious Tolerance. CATO Institute. https://www.cato.org/commentary/how-end-violence-middle-east-answer-religious-tolerance

Doxsee, C., Palmer, A., and McCabe, R. (2024). Chapter 6: Middle East. Global Terrorism Threat Assessment 2024, 71-91. Center for Strategic and International Studies (CSIS). http://www.jstor.org/stable/resrep57903.9

Gupta, A. K. (2015). THE RISE OF ISLAMIC STATE: CLASH WITHIN CIVILIZATION. The Indian Journal of Political Science, 76(3), 369–373. https://www.jstor.org/stable/26534848

Hahn, P. (2020). 9/11: causes and lingering consequences. The Ohio State University. https://www.osu.edu/impact/research-and-innovation/hahn-september-11

Jones, T. C. (2012). America, Oil, and War in the Middle East. Journal of American History, 99, 208-218. https://doi.org/10.1093/jahist/jas045

Kishi, K., and Theodorou, A. E. (2016). 6 facts about religious hostilities in the Middle East and North Africa. Pew Research Center. https://www.pewresearch.org/short-reads/2016/07/07/6-facts-about-religious-hostilities-in-the-middle-east-and-north-africa/

Lutz, B. J., and Lutz, J. M. (2015). Globalization and Terrorism in the Middle East. Perspectives on Terrorism, 9(5), 27–46. http://www.jstor.org/stable/26297432

Mahmood, R., and Jetter, M. (2019). Military Intervention Via Drone Strikes. Institute of Labor Economics. https://ssrn.com/abstract=3390307

Norlen, T. C. (2022). Middle East Pre‐Existing Conditions: Regional Security after Covid‐19. Middle East Policy, 29(1), 104–124. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9115253/

Rahman, K. (2010). US Middle East Policies and their Consequences. Policy Perspectives, 7(1), 33–60. http://www.jstor.org/stable/42909253

Spierings, N. (2019). The Multidimensional Impact of Islamic Religiosity on Ethno-religious Social Tolerance in the Middle East and North Africa. Social Forces, 97, 1693-1730. https://doi.org/10.1093/sf/soy092

Thomas, T., and Chaleff, I. (2017). Moral Courage and Intelligent Disobedience. InterAgency Journal, 8, 58-66. https://thesimonscenter.org/wp-content/uploads/2017/03/IAJ-8-1-Winter2017-pg58-66.pdf

Generational Disparities in Elderly-Care Expectations between Chinese Immigrants Who Arrived Before 1990 Versus Those Who Arrived after 2010

By Lin Zhang, Dr. Kah-Jin Jeffrey Kuan

Abstract

This study examines the differing expectations in regards to elderly care within the Chinese immigrant community[1] through two cohorts: immigrants who arrived before 1990 and those who arrived after 2010. Using existing research done in the field of cultural sociology, ethnic studies, and Confucian teachings, the study delivers an analytical research of how those factors combine to shape one's view on co-residence with elder parents, providing financial support, and institutionalized elder care. For Chinese immigrants who arrived before the 1990s, their goals generally fall under the following three categories: political refuge, escape from economic hardships, and escape from war. Prior to the 1990s, China's economy was largely agrarian, with a significant percent of its population laboring on farmlands. On the other hand, for Chinese immigrants who arrived after 2010, their reasons for immigration are as follows: pursuit of higher education, accumulation of wealth in China, and greater protection of property rights. The differences behind the motivation for immigration ultimately influenced their view on elderly care where the older generation of Chinese immigrants holds the traditional value of the collectivist culture and leans towards co-residence, while the younger generation of Chinese immigrants tend to hold a holistic view on the issue and exhibit more openness towards senior care facilities and direct financial support. The paper reveals the effect of time and the historical processes on cultural values, underscoring the importance of providing comprehensive elder care options within Chinese immigrant communities to suit the rapidly changing perspectives on the issue.

Introduction

The immigrant community in the United States is rapidly aging, with foreign-born elders making up 13.6% of the total elderly population of ages 65 and above (Mizoguchi, 2019). Although immigrants make up only about 14.3% of the U.S. population (American Immigration Council, 2023), the number of foreign-born elders is set to increase to 23.3% in the next four decades (Mizoguchi, 2019). The issue of elderly care is an undeniably growing concern within many immigrant communities, including that of the Chinese immigrant community. However, existing research and study done on eldercare expectation within the Chinese immigrant community exhibits a homogenous view on the otherwise culturally diverse community, failing to address the distinct needs of individual households that occurred as a result of generational disparities. The role that filial piety--a core Confucian value that underscores the importance of paying respect and care towards one's guardians (Kyeong, Yu, & Setoh, 2025)--plays in eldercare expectations, as well as the differing levels of acculturation experienced by Chinese immigrants over generations, are often overlooked. This, in turn, generates a gap in the understanding of the motivations behind choosing certain eldercare options among the culturally

diverse Chinese immigrant community, especially between those who arrived before the 1990s versus those who arrived after 2010, where filial obligation plays a heavier role in the former compared to the latter (Kim & Dong, 2019).

This paper aims to bridge the gap in understanding the motivation behind choosing a particular eldercare option by thoroughly examining the role of push and pull factors behind immigration across time periods. It argues that a strong correlation exists between the motivation behind immigration and the expectation on the issue of eldercare, which changed drastically across generations due to socioeconomic influences, exposure and openness towards western influences, and differing levels of acculturation. By examining literature that document an individuals' process of immigration, scholarly reports of relevant experiments, and theoretical frameworks, the paper reveals the factors that stem from motivations behind immigration and how those factors play a role in eldercare expectations.

The significance of this paper lies within its ability to bridge the gap in understanding the multifaceted beliefs within the Chinese immigrant community that was otherwise thought to be homogenous. Understanding one of the many nuanced perspectives of the Chinese immigrant community that is often overlooked (Zeng, Brasher, Gu, & Vaupel, 2018), the paper highlights the diversity that can exist in monoracial communities and addresses the need to cater eldercare towards the need of each distinct perspective.

Methodology

The paper takes on a theoretical approach in analyzing differing expectations of Chinese immigrants who arrived before 1990 to those after 2010 on the issue of eldercare. The research relies on literary works across both time periods, scholarly articles, and theoretical framework related to Chinese immigrants' eldercare practices. The paper aims to closely study the core values of the Confucian teaching, filial piety, acculturation theory, and generation theory in order to bridge the gap in the understanding of the change that took place on Chinese immigrants' expectations on eldercare.

The study examines how socioeconomic changes can generate diversity within a culture, zooming in on the particular issue of differing perspectives that can arise towards the option of family-based eldercare, which adheres to the value of filial piety, versus institutionalized eldercare, an alternative option that is gaining popularity due to a greater level of influence and openness towards western culture. The theoretical implications of this paper relies solely on peer-reviewed and publicly available material.

Literature Review

Existing frameworks on eldercare expectation within the Chinese immigrant community demonstrates the overarching influence of the Confucian teaching, filial-piety, where is it said that a positive correlation exists between the sense of filial obligation and the caregiver's burden of meeting tangible and intangible needs of the elder, and thus children who did not honor their filial obligations felt an increasing sense guilt (Xiao, 2024). In other words, current research

places a heavy emphasis on the culture of collectivism, filial obligations, and interdependence between family members on the issue of eldercare within Chinese immigrant communities (Guo, Lemke, & Dong, 2021). The framework is valid to the extent that it is largely applicable at explaining the motivation behind eldercare options of Chinese immigrant who arrived to the United States prior to 1990, while failing to capture the nuanced perspectives and expectations on eldercare that arose over the next couple of decades from Chinese immigrants who arrived after 2010 due to socioeconomic changes that took place in China.

Chinese immigration to the United States prior to 1990 can be examined in two waves: the 1850s and 1870s (Hooper & Batalova, 2015). Although a seemingly significant time disparity exists between the 1850s-1870s and 1990s, it is important to take into consideration that the Chinese Exclusion Act, which strictly prohibited Chinese laborers from entering the United States, was put into place in 1882 and was not repealed until 1943. Even so, for the next several decades following 1943, quotas were put in place to strictly control the number of Chinese immigrants who can come to the United States per year. Therefore, prior to the 1990s, the 1850s to 1870s was one of the few major waves of Chinese immigration that presents observable trends. During this period, political unrest and economic hardships left many Chinese people with no other options but to seek laborer work in western United States (Hooper & Batalova, 2015). Starting with the upheaval of the Taiping Rebellion, an uprising against the Manchu rulers that lasted for 14 years and resulted in over 20 million casualties (Britannia), it swept through China from the Yangtze River to Nanjing. The massive destruction and displacement of the power structure became a major push factor that drove Chinese people to the lands of the United States in order to escape the turmoil in their home country (Heffernan). Meanwhile, lured by the California Gold Rush of 1849 and the promise of the richness of the "Jinshan" (Gold mountains), Chinese migrants from Guangdong and Fujian began arriving in large numbers (William, 2023). Moreover, due to continuous debt brought on by generational poverty, many Chinese peasants left their homes to seek economic opportunities aboard--with a major one being the need for Railroad workers during the 1860s in the construction of the Transcontinental Railroad by the Pacific Central Railroad company (Chang, 2020). Although the influx of Chinese migration was ultimately cut short by the Chinese Exclusion Act of 1882 (Britannica), by the end of the 1870s, there were over 100,000 thousand Chinese immigrants living in the United States (U.S. Census Bureau). After the repeal of the Chinese Exclusion Act, strict quotas were still placed on the number of Chinese immigrants. It was not until the passage of the Immigration and Nationality Act in 1965 that more Chinese immigrants were able to come to seek refuge from the political instability in China. In 1989, university students in China who have been influenced by western teachings started a democratic reform movement. However, the movement was met with harsh response from the Chinese Communist Party (CCP). The Tiananmen Square Incident was one such student demonstration that was crushed by the government's use of military force (Britannica). As a result, the U.S. government responded by granting permanent resident status to over 50,000 Chinese students who were seeking safety under the Chinese Student Protection Act (United States Congress). In summary, political unrest

and economic hardships compelled Chinese immigrants to come to the United States to merely sustain life and to escape political persecution.

Brought together by similar backgrounds and experiences, the earliest Chinese immigrants formed tight communities that greatly provided convenience for its members in a foreign country, which later became known as Chinatowns (Zhou & Lee, 2013). Gradually, the establishment of Chinatowns attracted Chinese migrant workers and refugees, who typically hold a lower socioeconomic status, for the decades to come (Yamashita, 2013). Chinatowns uphold a strong sense of traditional value because one of the main reasons why migrants were attracted here is because it allows them to "engage their ancestral homelands via organizations and that organizational transnationalism contributes to strengthening the infrastructure and symbolic systems of the ethnic community" (Zhou & Lee, 2013). With one of the most prominent values in Chinese culture being filial piety, an experiment done on the Chinese elders (ages 60 and above) living in Chinatowns in the Greater Chicago area reports the participants have an overall high expectation in receiving filial care from their children (Dong, Zhang, & Simon, 2023). Despite how the study was done on later generations of Chinese descendants in Chinatown, their expectations are still reflective of that of the first generation of residents due to the lack of population movement caused by poverty and a low level of English proficiency (National Library of Medicine) in and out of the towns, keeping the community isolated from influences of other cultures. In summary, this exemplifies how Chinese immigrants who were driven to the United States out of the motivation to escape war, poverty, or prosecution, who also hold a lower socioeconomic status, have a tendency of choosing to reside in Chinatowns, where strong traditional values are upheld within the communities, one of which being the expectation to care for one's elders.

In comparison, Chinese immigration to the United States after 2010 are primarily driven by the pursuit of higher education and accumulation of wealth in China (Tian, Im, Mukherjee, & Budiman, 2024). China's access to the World Trade Organization in 2001 sparked an era of a towering level of export, which shaped China into a world trade titan by 2010 (UN Trade and Development). This period of intense economic growth in China significantly improved the quality of life for its citizens, where the percentage of people living in poverty decreased from 73% to 32% from 1990 to 2003 (World Trade Organization). Moreover, the average disposable income of households in China also increased drastically, as "just 4 percent of urban Chinese households[2] were in 2000--but 68 percent were in 2012" (Barton, Chen, & Jin, 2013). The acquisition of this new affluence prompted many Chinese to look for a promising future aboard, mainly in Canada and the United States. In 2010, 58% of immigrants in the EB-5 Immigrant Investor Program--a program that allows foreign investors to obtain a permanent residency in the country that they invest in--were Chinese nationals. Among the total number of immigrants in 2010, 34% of this group was able to obtain their legal status as immediate relatives of U.S. citizens (McCabe, 2012). In other words, after obtaining resident status, a large number of the Chinese immigrants also brought their parents over. The financial freedom that enabled many Chinese people to immigrate to the United States was often accompanied by the desire for better

educational resources. Families from a higher socioeconomic status exhibit a stronger desire towards the pursuit of self-development through the experience of studying abroad (Wang, 2024). With the arrival of the more affluent immigrants, new ethnic enclaves are formed in suburban areas rather than the traditional Chinatowns (Yamashita, 2013). Although still retaining many features that brought convenience for newly arrived Chinese immigrants, these ethnoburbs--more affluent, suburban areas (Tam 2024)--also promoted assimilation through the encouragement of its residents' social and civic participation, which became a critical site for the immigrants to engage with American culture (Nicolaides). In comparison to traditional Chinatowns, ethnoburbs exhibit a higher degree of racial diversity, for "in the City of Monterey Park, heart of the *ethnoburb* where the percentage of Chinese population was the highest among all ethnoburban communities, the Chinese accounted for 36% in 1990" (Bureau of the Census, 1990). "In Chinatown, however, ethnic Chinese was majority, accounting 62% of total population" (Seo 1992). Studies have also shown that the average age of its residents is significantly younger in inhabitants than Chinatowns, while the average English-proficiency level is higher in ethnoburbs than Chinatowns (Wei). The educational system in ethnoburbs embraces the western style of teaching, where the sense of individuality and self-consciousness is highly valued (Chen, 2024), straying away from the traditional Chinese value on filial piety with its strong emphasis on obedience and collectivism. This weakens the sense of filial obligation within the household because the younger generation are told to explore for their ideal lifestyles rather than doing something simply because they are told to, reflecting the American value of individualism. In summary, with the accumulated wealth that many households in ethnoburban areas already have and with their openness towards western influence, where filial care is no longer an obligation as households express more openness towards the idea that eldercare facilities can likely provide more professional health care (Cheung, 2020). Due to the constraints of work or education, many Chinese adults feel that they are unable to provide the best instrumental support to their elders, thus senior care facilities are viewed as a better alternative out of concern for the health of their parents (National Library of Medicine).

Although the existing framework did appropriately consider the role of filial piety in eldercare expectations among the Chinese immigrant community, making households within these communities lean towards the option of co-residence, it fails to capture the rise of another westernized perspective that came into the scene as a result of socioeconomic changes in China, both of which can be traced back to the immigrant's motivation for moving to the U.S. The socioeconomic change that took place was what drastically altered Chinese immigrants' motivation for immigration. Because of this change, very different push and pull factors can be observed for these two distinct waves of immigrants. The observed difference in expectations regarding eldercare between these two cohorts is the product of their distinct reasons behind immigration.

Conceptual Framework

The study draws from the core Confucian value of filial piety-- the idea of always paying respect and care towards one's guardians (Kyeong, Yu, & Setoh, 2025). Chinese immigrants who arrived prior to 1990 primarily lived in isolated ethnic enclaves known as Chinatowns, where much of the daily cultural practices are retained. Studies have shown that the immigrants before 1990, who were often driven by the motivation to escape from poverty, war, and famine, retained strong familial connections and traditional eldercare expectations (Chow, 2001). In contrast, immigrants who arrived after 2010 often gathered in ethnoburbs, where the ties to traditional values are significantly weaker while the individualism and institutionalized eldercare is much more accepted (Zhang & Montgomery, 2003). In summary, the change in the level of attachment with the Confucian value serves as an indicator to one's adherence to the traditional expectation of co-residence and family-based care.

The gradual detachment to the theory exhibited by the newer generation of immigrants can be explained and examined through the acculturation theory. Under Berry's Model of Accumulation, older immigrants who reside in Chinatowns exhibit separation from the host culture as their communities reject assimilation through its low English proficiency levels and the residents' desire to be around people whom they can relate to as a social minority in the United States facing social and economic discriminations (Wang). In comparison, Chinese immigrants who arrived after 2010 exemplify integration, for they typically exhibit a higher level of English proficiency and live in communities that encourage integration through political participation and westernized educational systems (Chen, 2024). As defined by Berry's model, although the newer generation of Chinese immigrants often exhibit a higher level of integration into American culture, they still maintain some traces of their culture of origin by often choosing to reside in ethnoburbs.

Another framework that this study draws from is the Cohort theory, the study of different populations in "cohort" of different time periods, regions, and life experiences (Nickerson, 2023). Chinese immigrants who arrived before 1990 can be categorized as one cohort due to their similar motivations behind immigration, while Chinese immigrants who arrived after 2010 can be categorized as one cohort for the same reason. Although immigrants who arrived before the 1990s encompassed a broad time range, these earlier immigrants all experienced periods of great political instability and economic hardship (Heffernan), which drove them to seek asylum in another country while holding on to most of their cultural identities, for collectivism is highly valued in ethnic Chinese culture. It wasn't until after 2010 that a significant shift in the overall motivation behind immigration was seen with a newer wave of immigrants. The newer generation of immigrants shared the same experience of the socioeconomic changes brought on by the industrial bloom of the 2000s in China, where a sense of liberalization was accompanied by the rise of consumer culture. The influences of time periods heavily influenced the life experiences of the immigrants, shaping their motivation behind immigration, resulting in differing expectations on many aspects of life.

The study proposes the Acculturative-Cohort Framework on Eldercare Expectation. This framework takes into consideration the time of immigration, the motivation behind immigration,

and the level of acculturation in examining the differing perspective on eldercare within Chinese immigrant communities. This framework allows a deeper understanding on the gradual shift of expectation on eldercare within Chinese immigrant communities, filling in the gap in the current understanding of these expectations, which are otherwise categorized monolithically. The framework presents the often overlooked nuance within one culture by accounting for the influences of historical context.

Discussion

The Acculturative-Cohort framework takes into account the influence of the time of immigration and the corresponding level of integration into western culture on one's expectation in eldercare. In contrast to existing models that view the Chinese immigrant community homogeneously, this framework considers the added nuance of acculturation and cohort effects in shaping one's point of view.

While existing frameworks on eldercare expectations in Chinese immigrant communities tend to underscore the critical role of filial piety, such assumptions can be overly broad, overlooking the socioeconomic changes in China that ultimately transformed the view of the newer generation of Chinese immigrant, which is observable through their drastically different motivation for immigration to the United States. This proposed framework is not to debunk existing research on eldercare expectations but to convey that these frameworks are only valid to the extent that the Chinese immigrant community has a monolithic and rigid culture, which is not the case; rather, the community, although sharing many common beliefs, has a plethora of differing perspectives, which can be captured through the Acculturative-Cohort Framework.

The framework bridges the gap in our understanding of the changing expectation on eldercare. It demonstrates where and how the gradual shifts took place rather than giving the appearance of a sudden disconnection. While both Berry's Model and the Cohort theory capture the reasons behind choosing a particular eldercare option for immigrants who arrived before 1990 and those who arrived after, the Acculturative-Cohort Model merges the two frameworks and incorporates the motivation behind immigration in context of socioeconomic changes to better understand how a shift of preferences took place. Berry's Model accurately asserts that acculturation can be broken down into categories, such as separation and integration. The Acculturative-Cohort Model supports this idea but goes a step further by explaining how people are placed in those categories. It adds that individuals are categorized based on shared responses to different historical processes. Additionally, while the proposed model agrees with the Cohort Theory that historical processes can serve as increments in categorizing individuals in groups, the key difference between Cohort Theory and the proposed model is how they define groupings. Cohort Theory groups individuals based on the shared experiences of the same historical events. However, this approach overlooks the fact that people from different backgrounds, such as varying socioeconomic statuses, may respond very differently to the same event. In contrast, the proposed model groups individuals based on the similar effects of different historical processes. For example, the Gold Rush and the construction of the Transcontinental Railroad were separate

events, but both incentivized poverty-stricken Chinese immigrants to come to the United States. The proposed model emphasizes these common effects rather than the specific events. By focusing on shared outcomes rather than shared events, this model reduces the risk of making inaccurate assumptions when categorizing people in groups. A possible flaw of the Acculturative-Cohort Model is that it might overgeneralize the impact of socioeconomic changes. With too much emphasis on categorizing immigrants based on specific events, the model can overlook distinct cases where immigrants are motivated to move due to factors that arise from events other than the major political, social, or cultural changes of their time, which causes the model to make overarching assumptions.

Conclusion

This study examines the differing expectations in regards to elderly care within the Chinese immigrant community through two cohorts: immigrants who arrived before 1990 and those who arrived after 2010. Through the exploration of literature work across both periods, scholarly articles, and related frameworks, the paper presents that the differences in expectations are influenced by socioeconomic changes in China, which can be observed through analyzing the motivation behind immigration. The difference in the motivations then translates into varying levels of openness towards western influences, which ultimately creates deviance within a culture that is otherwise thought to be homogenous.

Drawing from the existing Cohort theory and Berry's Model, the study presents how and why a gradual detachment from the Confucian value of filial piety occurred in the context of a changing social and economic environment in China. Going beyond the assumptions of tradition versus assimilation when it comes to observing the immigrant communities in the United States, the study presents a deeper understanding of the nuanced diversities that exist in between the very extremes of tradition and assimilation.

The value of the study lies within its ability to demonstrate the diversity behind a culture that is often thought to be monolithic. It also provides critical information for healthcare providers and law makers to better cater eldercare services towards the diverse needs of their communities.

Despite greater affluence and increased integration into mainstream American culture among newer generations of Chinese immigrants, systematic discrimination against the community still exists; it simply operates in a quieter, more insidious form today. Many Chinese immigrants still face limited social mobility; not because they lack wealth, but because they lack power as a racial minority. True mobility requires both. This paper calls for further investigation into how systemic racism continues to operate within the Chinese immigrant community, as well as whether the American culture, the so-called melting pot, is truly moving toward a community that embraces the pluralism of decoloniality.

Works Cited

Wang, Xi, Sarah L. Haft, and Qing Zhou. “Reasons for Migration, Post-Migration Sociocultural Characteristics, and Parenting Styles of Chinese American Immigrant Families.” Children (Basel), 24 Mar. 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC10136785/. Accessed 2 July 2025.

Tian, Ziyao, and Cecilia I. “Asian American Immigrants’ Experiences Adjusting to Life in the U.S.” Pew Research Center, 9 Oct. 2024, https://www.pewresearch.org/race-and-ethnicity/2024/10/09/asian-american-immigrants-experiences-adjusting-to-life-in-the-u-s. Accessed 4 July 2025.

“Cultural Orientation in China: Differences Across Five Generations of Employees.” ResearchGate, https://www.researchgate.net/publication/303470542_Cultural_Orientation_in_China_Differences_Across_Five_Generations_of_Employees. Accessed 6 July 2025.

“The Older Foreign-Born Population in the United States.” U.S. Census Bureau, https://www.census.gov/content/dam/Census/library/publications/2019/acs/acs-42.pdf. Accessed 8 July 2025.

“U.S. Immigration Statistics.” American Immigration Council, https://map.americanimmigrationcouncil.org/locations/national/. Accessed 10 July 2025.

Wiley. “Article Title Not Provided.” AGS Journals, https://agsjournals.onlinelibrary.wiley.com/doi/am-pdf/10.1111/jgs.17511. Accessed 12 July 2025.

Zeng, Yi, et al. “Older Parents Benefit More in Health Outcome from Daughters’ than Sons’ Emotional Care in China.” Journal of Aging and Health, Dec. 2016, https://pmc.ncbi.nlm.nih.gov/articles/PMC5947966/. Accessed 14 July 2025.

Kyeong, Yena, et al. “Filial Piety Across Sociocultural Context and the Life Span.” Current Directions in Psychological Science, 2025, https://journals.sagepub.com/doi/full/10.1177/09637214241312630. Accessed 16 July 2025.

Xiao, Chun, et al. “Filial Piety and Older Adult Caregiving among Chinese and Chinese-American Families in the United States: A Concept Analysis.” BMC Nursing, 13 Feb. 2024, https://bmcnurs.biomedcentral.com/articles/10.1186/s12912-024-01789-0. Accessed 18 July 2025.

Guo, Min, Amy Lemke, and XinQi Dong. “Sources of Intergenerational Conflict in Chinese Immigrant Families in the United States.” Journal of Family Issues, Sept. 2022, https://pmc.ncbi.nlm.nih.gov/articles/PMC9512278/. Accessed 20 July 2025.

“Chinese Immigrants in the United States.” Migration Policy Institute, https://www.migrationpolicy.org/article/chinese-immigrants-united-states-2013. Accessed 22 July 2025.

Encyclopædia Britannica, Inc. “Taiping Rebellion.” Encyclopædia Britannica, https://www.britannica.com/event/Taiping-Rebellion. Accessed 24 July 2025.

"Http://journals.sagepub.com/doi/abs/10.1177/0887302x07303626 | Request PDF." ResearchGate, https://www.researchgate.net/publication/328039672_httpjournalssagepubcomdoiabs1011770887302X07303626. Accessed 26 July 2025.

"California Gold Rush and Immigration." EBSCO Information Services, https://www.ebsco.com/research-starters/history/california-gold-rush-and-immigration. Accessed 28 July 2025.

Yamashita, Kohei. "A Comparative Study of Chinatowns Around the World: Focusing on the Increase in New Chinese Immigrants and Formation of New Chinatowns." Japanese Journal of Human Geography, 26 Jan. 2018, https://www.jstage.jst.go.jp/article/jjhg/65/6/65_527/_article. Accessed 30 July 2025.

"Transnationalism and Community Building: Chinese Immigrant Organizations in the United States." ResearchGate, https://www.researchgate.net/publication/258125517_Transnationalism_and_Community_Building_Chinese_Immigrant_Organizations_in_the_United_States. Accessed 1 July 2025.

Barton, David, Yan Chen, and Amy Jin. "Mapping China's Middle Class." McKinsey & Company, 1 June 2013, https://www.mckinsey.com/industries/retail/our-insights/mapping-chinas-middle-class. Accessed 3 July 2025.

"Http://journals.sagepub.com/doi/abs/10.1177/0887302x07303626 | Request PDF." ResearchGate, https://www.researchgate.net/publication/328039672_httpjournalssagepubcomdoiabs1011770887302X07303626. Accessed 5 July 2025.

"Suburban Metamorphosis." Center for the Study of Women, 29 June 2023, https://csw.ucla.edu/2016/12/19/suburban-metamorphosis. Accessed 7 July 2025.

Tam, Chi-Chung. "Ethnic Enclaves and Ethnoburbs: Are There Differences in Associations with Juvenile Offense Type Among Asian Americans?" Journal of Ethnic & Cultural Diversity in Social Work, 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC9782721/. Accessed 9 July 2025.

"Psychological Well-Being of Older Chinese Immigrants Living in Australia: A Comparison with Older Caucasians." ResearchGate, https://www.researchgate.net/publication/305074964_Psychological_well-being_of_older_Chinese_immigrants_living_in_Australia_a_comparison_with_older_Caucasians. Accessed 11 July 2025.

"Chinese Americans: History, Modern Era, History of Chinese Immigration, Settlement Patterns." EveryCulture.com, https://www.everyculture.com/multi/Bu-Dr/Chinese-Americans.html. Accessed 13 July 2025.

"Chinese Immigrants in the United States." Migration Policy Institute, https://www.migrationpolicy.org/article/chinese-immigrants-united-states-2010. Accessed 15 July 2025.

"S.1216 - 102nd Congress (1991-1992): Chinese Student Protection Act of 1992." Congress.gov, https://www.congress.gov/bill/102nd-congress/senate-bill/1216. Accessed 17 July 2025.

Integrating Graph Neural Networks with Protein Language Models for Functional Annotation By Sreyas Prabu

Abstract

Proteins drive nearly every biological process. Their ability to function depends on their amino acid sequences and the complex ways those sequences fold into three-dimensional structures. Researchers have studied these two features separately for years, either by modeling shapes or analyzing sequence patterns. Recent advances in AI are changing that. Graph neural networks (GNNs) can map protein structures as spatial graphs, whereas protein language models (PLMs), such as ProtBERT, learn from millions of sequences (Brandes et al., 2022; Rives et al., 2021). These methods could be improved and optimized for use in other fields. This paper examines how each method operates, what they offer individually, and how combining them enables a more accurate prediction of protein function.

Introduction

Automated protein function prediction is a central problem in bioinformatics, driven by the vast sequence databases with millions of entries. For example, while databases like UniProt contain on the order of 200 million protein sequences, only a few million have reliable functional labels. Traditional homology-based methods often fail to identify remote homologs and novel folds, motivating deep learning approaches that can generalize beyond simple sequence similarity. In parallel, the advent of protein language models (PLMs) trained on massive sequence corpora has provided powerful sequence representations. Models such as ProtBERT (Brandes et al., 2022), ESM (Rives et al., 2021), and others learn contextual embeddings of amino acid sequences that capture evolutionary and biophysical patterns (e.g., conserved motifs and domains) even in the absence of labeled data. These embeddings have shown great performance on diverse protein predictions, from secondary structure to subcellular localization.

Beyond DNA sequence, protein function is inherently tied to 3D structure and physical interactions. Graph neural networks (GNNs), and in particular graph convolutional networks (GCNs), are well-suited to model molecular structures by treating residues as nodes and their physical contacts as edges (Kipf & Welling, 2016). In recent work, Gligorijević et al. demonstrated that a two-stage architecture, using a pre-trained LSTM language model and a multi-layer GCN, significantly outperforms sequence-only or structure-only baselines Gligorijević et al., 2021). In particular, integrating sequence embeddings and structural graph convolutions resulted in significantly improved function prediction accuracy compared to using contact maps with one-hot encoding (Gligorijević et al., 2021). These results highlight the complementary nature of sequence and structure information: rich embeddings from PLMs encode long-range evolutionary context, while GCNs capture spatially localized interactions. In this paper, we extend the hybrid approach by utilizing a modern transformer-based PLM (ProtBERT) and an explicit GCN architecture to determine the optimal method for predicting gene ontology (GO) annotations.

Our contributions include:

- Integrated architecture: We detail a combined model that inputs ProtBERT embeddings of the amino acid sequence and a GCN on the protein's residue contact graph, merging their outputs for GO-term prediction.
- Code-level design: We provide pseudocode and implementation notes for data preprocessing (Protein Data Bank (PDB) parsing, graph construction) and network layers.
- Data pipeline: We propose an automated system for expanding the training dataset by retrieving new PDB structures and parsing individual chains into graphs, improving diversity and reducing noise.
- Experimental results: We report initial training outcomes (loss curves over 10 epochs) that demonstrate effective learning behavior.
- Implications: We reflect on how this approach could scale to proteome-wide annotation and aid drug discovery by identifying functional sites.

Methods

Data Preparation and Preprocessing:

Our training data consisted of proteins with known functions (Gene Ontology annotations) and available 3D structures. We preprocessed data using the following steps:

- Automated structure retrieval: We periodically queried the RCSB PDB repository to obtain new or updated entries. For each PDB ID, we downloaded the mmCIF file. Using the Biopython PDB module, we parsed the file into a Structure object, which conveniently handled atomic coordinates and chain segmentation (Cock et al., 2009).
- Chain-specific parsing: Multi-chain PDB entries were split, with each chain treated as a separate protein instance. Using the Structure object, we iterated over chains and extracted only the coordinates of that chain's residues. This process isolated the geometry and sequence of each polypeptide, avoiding the mixing of interfaces (unless interface information was desired).
- Residue graph construction: For each chain, we built an undirected graph where each node corresponded to a residue (typically represented by the Cα atom or centroid). We then added an edge (i, j) if the distance between residues i and j fell below a threshold of 10 angstroms, indicating a contact in 3D space. This yielded an adjacency matrix. This residue-level graph encoded the protein's folded geometry.
- Sequence embedding extraction: We obtained the amino acid sequence of the chain using the Hugging Face library, and we loaded the pre-trained ProtBERT model (Rostlab's version), which provided 768-dimensional embeddings for each residue token (Brandes et al., 2022; Rives et al., 2021). The raw sequence was tokenized and fed into ProtBERT; we used the last hidden state as the sequence feature tensor. These embeddings

summarized context over the entire protein, incorporating information learned from hundreds of millions of sequences.

- Function labels (GO terms): For supervised learning, each protein chain was associated with a set of GO annotations (molecular function, biological process, etc.) obtained from curated databases. We treated function prediction as a multi-label classification: the model outputs a probability for each of the GO terms. During training, we used the binary cross-entropy loss across terms.

Model Architecture

Our network had two main branches that processed the protein sequence and structure separately, and then merged:

Sequence branch (ProtBERT): We embedded the raw amino acid sequence with the pre-trained ProtBERT model. Formally, given input tokens, the model outputs embeddings. We froze ProtBERT weights so they served as a static feature extractor, reducing computational cost and overfitting. These embeddings were then passed through a linear layer to project to the GCN input dimension.

Structure branch (GCN): We constructed a GCN operating on the residue contact graph. Each node, i, initially carried the ProtBERT embedding. We applied graph convolutional layers according to the method proposed by Kipf–Welling (Kipf & Welling, 2016). Each layer computes:

$$H^{(l+1)} = \mathrm{ReLU}(\tilde{D}^{-1/2} \cdot \tilde{A} \cdot \tilde{D}^{-1/2} \cdot H^{(l)} \cdot W^{(l)})$$

Where:

$H^{(l)}$: Node feature matrix at layer l

$W^{(l)}$: Trainable weight matrix at layer l

$\tilde{A} = A + I$: Adjacency matrix with added self-loops

$\tilde{D}$: Degree matrix of $\tilde{A}$

$\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2}$: Symmetric normalization of the adjacency matrix

ReLU: Rectified Linear Unit activation function

After several layers, each node's feature aggregated information from its spatial neighbors, capturing local 3D context.

Global pooling and output:

To create a representation of the entire protein, we applied global mean pooling over the final layer of node features. This produced a fixed-size vector regardless of the protein sequence length. If the output of the last GCN layer was H^(L_g) ∈ ℝ^(L × h), where L is the number of residues and h is the hidden size, then the mean-pooled vector was:

$$h_{\text{mean}} = \frac{1}{L} \times \sum_{i=1}^{L} H_{i,:}^{(L_g)} \in \mathbb{R}^h$$

This vector could be optionally concatenated with global sequence-level features. The resulting vector was passed through one or more fully connected (dense) layers with ReLU activations.

The final layer outputs K logits, each representing a GO term. These logits were passed through a sigmoid function to produce multi-label probabilities. In practice, we sometimes used more GCN layers or hidden units. We also experimented with variants such as Graph Attention (GAT) and multi-graph channels, but the above architecture sufficed for our initial tests. Importantly, by sharing the ProtBERT embeddings in both branches (i.e., using the same residue features for sequence and structure processing), we ensured that the network jointly refined how sequence context and spatial context contributed to function. This two-branch design was inspired by DeepFRI's two-stage model (Gligorijević et al., 2021), although we used a transformer-based PLM and an explicit graph pooling layer.

Training pipeline

We trained the network on proteins with known GO annotations in a supervised multi-label setting. Key training details included:

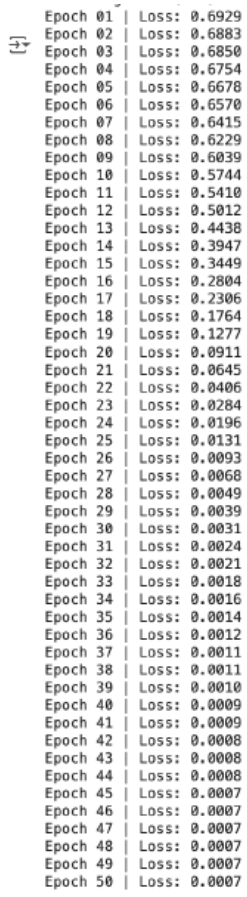
- **Loss function:** We used multi-label binary cross-entropy (BCE) across all Gene Ontology (GO) terms. This loss function was appropriate because each protein could be associated with multiple GO terms independently.
- **Optimizer**: We employed the Adam optimizer. Only the GCN and classification weights were learned.
- **Batching**: Graph inputs were batched using PyTorch Geometric, which handled varying graph sizes. Since proteins had different lengths, we padded sequences to the same token length in each batch for ProtBERT.
- **Epochs and early stopping:** We trained for approximately 10–20 epochs, monitoring training and validation loss. Early stopping was applied if validation loss plateaued. In our initial experiments, 10 epochs yielded substantial loss reduction. We also experimented with dropout (0.1–0.2) in the fully connected layers to prevent overfitting.
- **Data splits:** The dataset was split purely at the protein level, ensuring no overlapping sequences. We also ensured that proteins in different splits shared no high sequence identity (e.g., <30%) to test generalization (Radivojac et al., 2013).
- **Implementation**: Our code was implemented in PyTorch with PyTorch Geometric. Preprocessing (PDB parsing, graph creation) was performed in Python using Biopython (Cock et al., 2009). Sequence tokenization and ProtBERT inference used Hugging Face's Transformers library.

- **Benchmarks**: To benchmark performance, we tracked the model's loss through 50 epochs and compared the results.

Results

We evaluated the effectiveness of our models using the binary cross-entropy loss across epochs. Two primary configurations were tested: a standard Graph Convolutional Network (GCN) trained solely on protein structure data, and a hybrid GCN + PLM model that incorporates both structural and sequence-level information via protein language model embeddings.

The plain GCN model, trained using only node embeddings derived from ProtBert token IDs and structural connectivity from C-alpha distances, showed a steady and prolonged convergence pattern. Starting with an initial loss of 0.6929 at epoch 1, the model gradually reduced the loss over time, reaching 0.5744 by epoch 10 and 0.0911 by epoch 20. The convergence continued into deeper training, with losses dropping to 0.0131 at epoch 25 and 0.0007 by epoch 50. This slow but consistent decline in loss suggests the GCN was capable of learning useful structural patterns, though it required many epochs to reach optimal performance.



```
Epoch 01 | Loss: 0.6929
Epoch 02 | Loss: 0.6883
Epoch 03 | Loss: 0.6850
Epoch 04 | Loss: 0.6754
Epoch 05 | Loss: 0.6678
Epoch 06 | Loss: 0.6570
Epoch 07 | Loss: 0.6415
Epoch 08 | Loss: 0.6229
Epoch 09 | Loss: 0.6039
Epoch 10 | Loss: 0.5744
Epoch 11 | Loss: 0.5410
Epoch 12 | Loss: 0.5012
Epoch 13 | Loss: 0.4438
Epoch 14 | Loss: 0.3947
Epoch 15 | Loss: 0.3449
Epoch 16 | Loss: 0.2804
Epoch 17 | Loss: 0.2306
Epoch 18 | Loss: 0.1764
Epoch 19 | Loss: 0.1277
Epoch 20 | Loss: 0.0911
Epoch 21 | Loss: 0.0645
Epoch 22 | Loss: 0.0406
Epoch 23 | Loss: 0.0284
Epoch 24 | Loss: 0.0196
Epoch 25 | Loss: 0.0131
Epoch 26 | Loss: 0.0093
Epoch 27 | Loss: 0.0068
Epoch 28 | Loss: 0.0049
Epoch 29 | Loss: 0.0039
Epoch 30 | Loss: 0.0031
Epoch 31 | Loss: 0.0024
Epoch 32 | Loss: 0.0021
Epoch 33 | Loss: 0.0018
Epoch 34 | Loss: 0.0016
Epoch 35 | Loss: 0.0014
Epoch 36 | Loss: 0.0012
Epoch 37 | Loss: 0.0011
Epoch 38 | Loss: 0.0011
Epoch 39 | Loss: 0.0010
Epoch 40 | Loss: 0.0009
Epoch 41 | Loss: 0.0009
Epoch 42 | Loss: 0.0008
Epoch 43 | Loss: 0.0008
Epoch 44 | Loss: 0.0008
Epoch 45 | Loss: 0.0007
Epoch 46 | Loss: 0.0007
Epoch 47 | Loss: 0.0007
Epoch 48 | Loss: 0.0007
Epoch 49 | Loss: 0.0007
Epoch 50 | Loss: 0.0007
```

By contrast, the GCN + PLM model, which fused structural data with contextualized embeddings from ProtBert, demonstrated a far more rapid convergence. The model began with a

loss of 0.6673 at epoch 1 and reached 0.1725 by just epoch 3. Remarkably, by epoch 5, the loss dropped to 0.0005, and by epoch 10, it stabilized at 0.0001, indicating that the model had already achieved near-optimal classification performance. This dramatic improvement in both speed and accuracy highlights the significant value of incorporating protein language models into the representation learning pipeline.

```
Epoch 01 | Loss: 0.6673
Epoch 02 | Loss: 0.4916
Epoch 03 | Loss: 0.1725
Epoch 04 | Loss: 0.0133
Epoch 05 | Loss: 0.0005
Epoch 06 | Loss: 0.0002
Epoch 07 | Loss: 0.0002
Epoch 08 | Loss: 0.0001
Epoch 09 | Loss: 0.0001
Epoch 10 | Loss: 0.0001
```

These findings clearly show that the integration of pre-trained language model embeddings not only accelerates convergence but also enhances the predictive capability of the model. The GCN + PLM approach learns rich protein representations in far fewer epochs and achieves significantly lower loss values, underscoring the benefits of multimodal input fusion in protein function prediction (Yang et al., 2024).

Discussion

Our integrated GCN+PLM approach effectively combines sequence and structure information. The experimental trend of steady loss decline suggests the network successfully exploits both sources. DeepFRI's results similarly showed that jointly using graph convolutions and language-model features outperforms using either alone or with an LSTM (Gligorijević et al., 2021). Intuitively, the PLM (ProtBERT) captures long-range sequence patterns that may hint at function (e.g., motifs conserved in enzyme families), while the GCN enforces that spatially proximate residues can jointly influence function (e.g., a catalytic triad).

In addition, our architecture is flexible. For example, one could replace ProtBERT with other PLMs (ESM, ProteinBERT, or even fine-tuned variants) to use for more specific types of proteins. Our work demonstrates proof-of-principle, but more advanced GNN layers (graph attention, relational graphs) could be incorporated (Xu et al., 2019), increasing the efficiency of the model.

The proposed data expansion pipeline is a key future direction. By continuously retrieving new PDBs, splitting chains, and incorporating even predicted structures such as AlphaFold DB (Jumper et al., 2021), we can greatly enlarge the training set. Expanding the

quantity in which we train the model should improve generalization: training on a few thousand proteins risks bias, whereas including tens or hundreds of thousands would cover more of the sequence-structure space. DeepFRI authors similarly augmented their training with homology models to boost function coverage (Gligorijević et al., 2021). Chain-specific parsing increases sample count and ensures that each polypeptide's function is learned in context (e.g., avoiding cross-talk from multi-chain complexes).

One challenge in this pipeline is label noise: not all PDB chains have complete or accurate GO annotations. We may use thresholding or label cleaning, by excluding very general GO terms, to mitigate this. As the dataset grows, techniques like curriculum learning or self-training, where high-confidence predictions are added as pseudo-labels, could further improve robustness.

Beyond GO terms, our framework could be adapted to related tasks. The residue-level interpretability has the potential to highlight drug-binding sites or allosteric regions, aiding drug discovery efforts. Moreover, since ProtBERT and similar models can process extremely long sequences, our method is capable of handling large proteins or multi-domain chains, unlike older LSTM-based approaches, allowing it to decipher a larger array of naturally occurring proteins.

Limitations

Our current implementation is computationally heavy due to ProtBERT inference. In production, one could use distilled or smaller PLMs (Brandes et al., 2022) or freeze embeddings as we did. Also, 3D structure input means the method is limited by available models. However, with resources like AlphaFold DB providing predicted structures for nearly all sequences, this constraint is loosening (Jumper et al., 2021). Our automated pipeline could seamlessly ingest AlphaFold predictions in addition to experimental PDB entries.

Conclusion

We have presented a comprehensive framework for functional annotation of proteins by integrating graph convolutional neural networks with protein language model embeddings. Through a detailed description of the data preprocessing, architecture, and training pipeline, we demonstrate how to construct such a hybrid model in practice. Our initial experiments show promising learning behavior, and we expect that scaling up data via automated PDB retrieval and chain parsing will further enhance performance.

This work has significant implications. A scalable, accurate function predictor can accelerate large-scale genome and metagenome annotation, filling gaps left by homology methods. By pinpointing functional residues via integrated sequence-structure signals, the model could guide mutagenesis experiments or identify druggable sites. In drug discovery, annotating novel proteins (e.g., microbial enzymes or orphan GPCRs) can suggest new therapeutic targets or enzymes for bioengineering.

Works Cited

Brandes, N., Ofer, D., Peleg, Y., Rappoport, N., & Linial, M. (2022). ProteinBERT: A universal deep-learning model of protein sequence and function. Bioinformatics, 38(8), 2102–2110. https://doi.org/10.1093/bioinformatics/btac061

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. L. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics, 25(11), 1422–1423. https://doi.org/10.1093/bioinformatics/btp163

Gligorijević, V., Renfrew, P. D., Kosciolek, T., Leman, J. K., Berenberg, D., Vatanen, T., Chandler, C., Taylor, B. C., Fisk, I. M., Vlamakis, H., Xavier, R. J., Knight, R., Cho, K., Bonneau, R., & Sander, C. (2021). Structure-based protein function prediction using graph convolutional networks. Nature Communications, 12, 3168. https://doi.org/10.1038/s41467-021-23303-9

Jiang, Y., Oron, T. R., Clark, W. T., Bankapur, A. R., D'Andrea, D., Lepore, R., Funk, C. S., Kahanda, I., Verspoor, K. M., Ben-Hur, A., Koo, D. C. E., Penfold-Brown, D., Shasha, D., Youngs, N., Bonneau, R., Linial, M., Radivojac, P., Friedberg, I., & Mooney, S. D. (2016). An expanded evaluation of protein function prediction methods shows an improvement in accuracy. Genome Biology, 17, 184. https://doi.org/10.1186/s13059-016-1037-6

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. arXiv. https://arxiv.org/abs/1609.02907

Radivojac, P., Clark, W. T., Oron, T. R., Schnoes, A. M., Wittkop, T., Sokolov, A., Graim, K., Funk, C., Verspoor, K., Ben-Hur, A., Pandey, G., Yunes, J. M., Talwalkar, A., Repo, S., Jones, D. T., Friedberg, I., Bonneau, R., Bradley, P., Rex, D. K., ... Mooney, S. D. (2013). A large-scale evaluation of computational protein function prediction. Nature Methods, 10, 221–227. https://doi.org/10.1038/nmeth.2340

Rives, A., Meier, J., Sercu, T., Goyal, S., Lin, Z., Liu, J., Guo, D., Ott, M., Zitnick, C. L., Ma, J., & Fergus, R. (2021). Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. Proceedings of the National Academy of Sciences, 118(15), e2016239118. https://doi.org/10.1073/pnas.2016239118

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2019). How powerful are graph neural networks? arXiv. https://arxiv.org/abs/1810.00826

Yang, Y., Zhu, Q., Zhang, L., Li, M., Wang, J., & Peng, J. (2024). Protein function prediction with multitask deep neural networks. Nature Communications, 15, 1113. https://doi.org/10.1038/s41467-024-45089-3

BurstDetector: Using Machine Learning to Predict Starbursts in Galaxy Mergers

By Sakethram Badri and Vishesh Mishra

Abstract

When two galaxies merge together, several phenomena such as starbursts can ensue. Studying such phenomena is crucial to understanding how galaxies and stars interact, with the analysis of star formation rates providing insight into the many mysteries of our universe. Current methods of direct imaging and spectral analysis in analyzing mergers are mostly manual and not automated, oftentimes being prone to error as well. With large datasets emerging online through more and more readily available simulation data, more efficient methods must be developed to study such data. Machine learning techniques can expedite such processes, with this paper aiming to evaluate 3 techniques known for successful image classification in their success with automating the analysis of images from these datasets: Convolutional Neural Networks (CNN), Random Forest Algorithms (RF), and Support Vector Machines (SVM). Trained and cross-validated on image data from the Sloan Digital Sky Survey, our CNN ”BurstDetector” yielded the most success with an accuracy of 92.7% in detecting the occurrence of starbursts, demonstrating that CNNs tend to experience the most success in this image classification task of the 3 models evaluated. BurstDetector can also be run on a multitude of computers regardless of their GPU, making it computationally efficient. A computationally efficient model like BurstDetector is essential to being able to interpret the tremendous amount of data online. The study of the resultant stars forming in merging galaxies through their images can be pivotal to making new discoveries in the field of physics and astronomy, opening the door to revelations in the structure of the universe and even progress with dark matter.

Introduction

Galaxy mergers occur when two roughly equally sized galaxies collide against one another, resulting in a violent interaction between the contents of the two cosmic bodies. These events can cause phenomena such as intense bursts of star formation, known as starbursts. Understanding such phenomena is important for advancing our knowledge of galaxy evolution and the processes that shape the universe. Predicting when and where starbursts occur in merging galaxies could potentially provide insight into the mechanisms that trigger intense periods of star formation in the universe. Research on the activity of starburst mergers can even reveal the unseen distribution of dark matter within galaxies, which is one of the biggest mysteries in astrophysics.

Rationale

Having an accurate and efficient method of knowing which galaxies to study further will save time when analyzing mergers. With most current methods on conducting such analysis being manual, doing this research is extremely time consuming. Machine learning can

enable the efficient analysis of vast astronomical datasets, which in recent times have begun to develop more and more, and uses algorithms to predict phenomena with greater accuracy and speed. Machine learning techniques can revolutionize the way scientists interpret interactions between galaxies, providing predictive power that was previously unavailable through manual techniques. The scientific impact of the project extends beyond galaxy evolution, contributing to the field of astronomy by accelerating discoveries and improving data analysis techniques. Moreover, the predictive power of machine learning can transfer over to various other fields, ultimately advancing scientific research processes and technological innovation in society.

Data Collection

Firstly, we constructed an SQL query to filter for galaxy mergers and gather images of galaxy mergers in FITS format with their corresponding star formation rates from the Sloan Digital Sky Survey (SDSS) SQL image query [1] [2]. Given the limited availability of high-quality merger images, TensorFlow-based data augmentation techniques [3] were utilized to artificially expand the dataset while preserving astrophysical accuracy. 14517 images were created, and were split into 11613 training images and 2904 cross validation images (80/20 split).

- Geometric Transformations (Rotating, Flipping, Scaling)
- Color Transformations (Brightness, Saturation Adjustment)
- Addition of Artificial Noise (Blur, Sharpness)

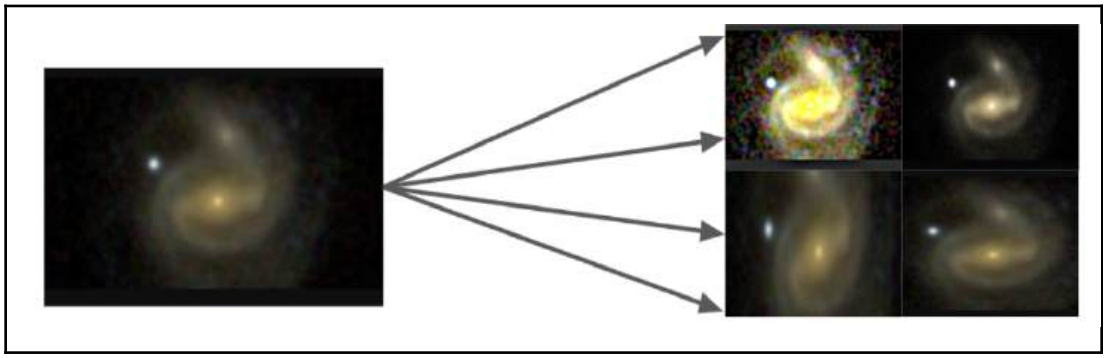


Figure 1: Example of how Tensorflow augments images to expand a dataset

Model Development

Three models with three different machine learning algorithms were trained and developed: Convolutional Neural Networks (CNN), Random Forest (RF), and Support Vector Machine (SVM) [4] [5]. Each model was programmed and executed in VSCode Jupyter Notebooks to implement and fine-tune the training process. To determine the most effective starburst classification model, the cross-validation dataset was fed through all three models,

and their performances were assessed using multiple evaluation metrics, including accuracy, precision, sensitivity, specificity, and F1-Score. After initial evaluations, we refined the most effective model, the CNN, by tuning parameters such as learning rates and improved preprocessing techniques to enhance key astrophysical features, such as gas density and tidal tails.

Results

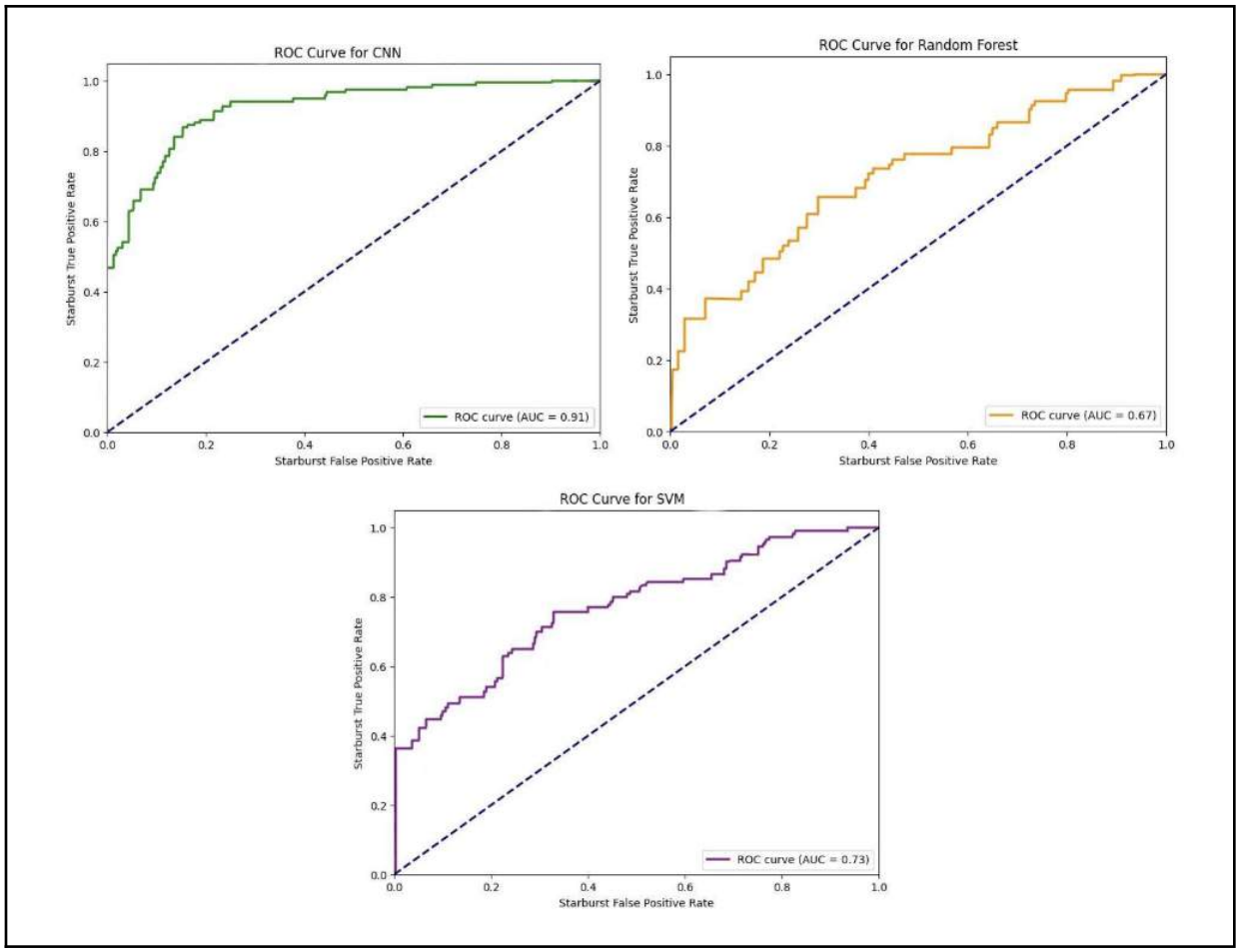


Figure 2: ROC curves comparing AUC across starburst classification models

As seen in Figure 2, the Convolutional Neural Network (CNN) yielded the most effectiveness with the highest AUC (Area Under Curve) of 0.91 on the ROC curves. A higher AUC signifies a more effective model in making accurate predictions.

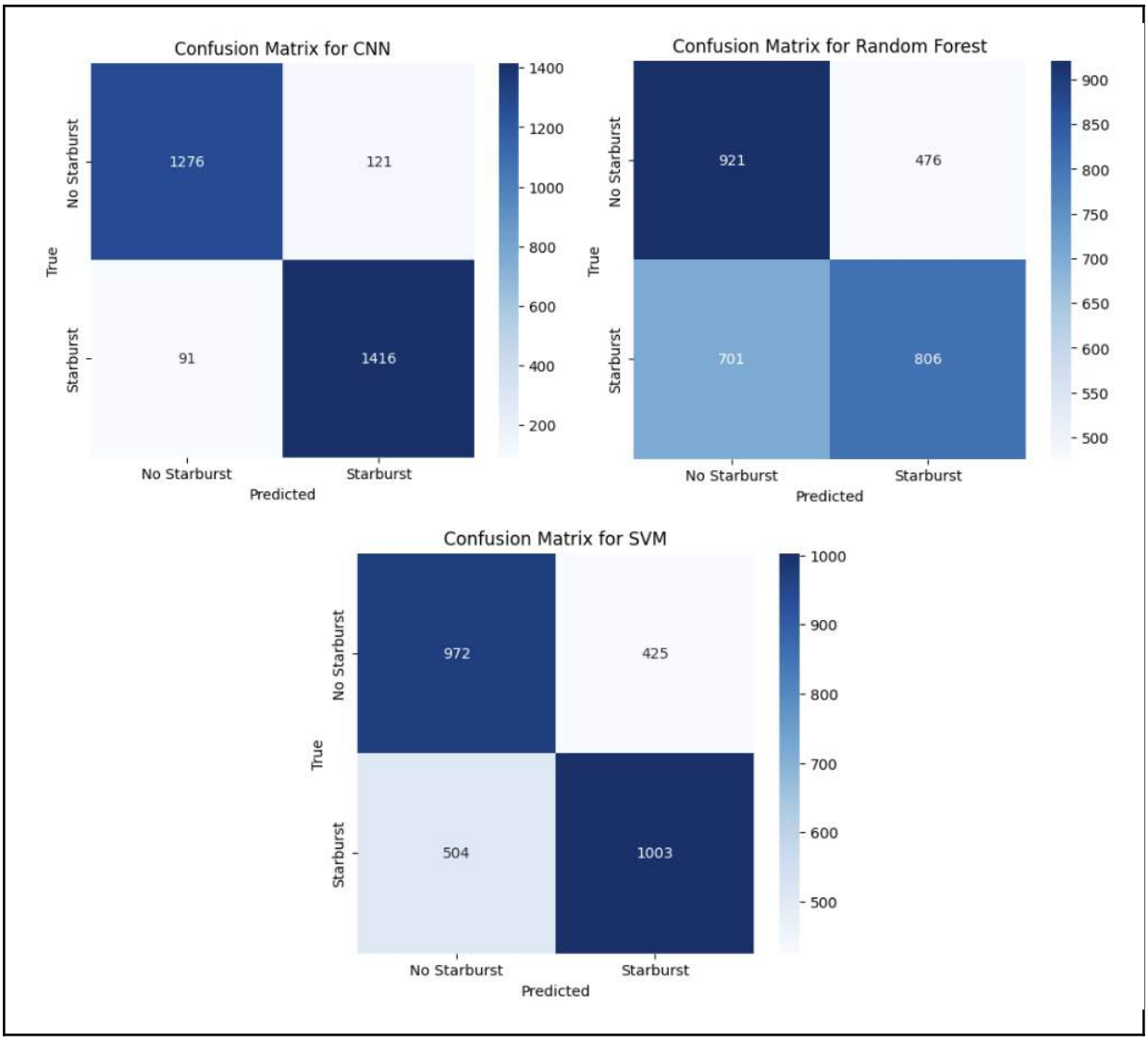


Figure 3: Confusion Matrices visually representing true and false positives and negatives

The Confusion Matrix further confirmed the CNN's success, showing the largest proportion of correct predictions, minimizing Type I and Type II errors.

CNN Metrics	Value
Accuracy	0.927
Precision	0.921
Recall (Sensitivity)	0.939
F1 Score	0.931
Specificity	0.913

RF Metrics	Value
Accuracy	0.595
Precision	0.629
Recall (Sensitivity)	0.535
F1 Score	0.578
Specificity	0.659

SVM Metrics	Value
Accuracy	0.681
Precision	0.702
Recall (Sensitivity)	0.666
F1 Score	0.684
Specificity	0.696

Figure 4: Metrics for each classification model. The CNN outperforms both of the other models in every metric.

Additionally, the CNN model achieved an overall accuracy of 92.7% in detecting starbursts, indicating its high reliability and effectiveness in identifying the phenomena of starbursts. These observations as well as the direct comparisons in precision, recall, accuracy, F1-score, sensitivity, and specificity between the 3 techniques make it clear that the Convolutional Neural Network was the most effective, and underscores the potential of machine learning techniques to improve astronomical predictions.

The Convolutional Neural Network (CNN) used multiple layers to progressively learn and extract important features from the galaxy merger images. The strength of CNNs lies in their ability to automatically learn hierarchical features from raw image data, without the need for manual feature extraction [5].

- 1. Edge Detection and Low-Level Features: The initial convolutional layers were responsible for detecting basic visual features such as edges and textures. Filters were applied to identify simple structures like gradients and pixel intensity, which allows the model to detect the boundaries between merging galaxies.
- 2. Shape Recognition and Mid-Level Features: As the image data passed through subsequent convolutional layers, the network recognized more complex patterns, such as shapes and spatial relationships. These features include the identification of spiral arms, elliptical galaxy shapes, and merging tidal tails.
- 3. Gas Density and High-Level Features: The deeper layers of the network were critical in

detecting astrophysical phenomena such as gas density. By analyzing regions with high gas
concentration and star formation activity, the model recognized features of starburst events, where star formation occurs due to gravitational interactions.

- 4. Starburst Classification: After extracting the layered features (edges, shapes, gas density) and combining the respective SFR data, the model used a fully connected layer to combine the learned information into a final classification. The final classification decision was based on the recognition of specific astrophysical features like tidal tails, gas density variations, and the morphology of interacting galaxies that typically precede starburst events

Processing each image in just 23 milliseconds, the CNN algorithm proves to be significantly faster than traditional human manual analysis. Such swift classification not only makes the process more efficient, but also allows for the rapid analysis of vast datasets. The resultant combination of high accuracy, speed, and ability to handle large datasets positions the CNN as a highly effective tool for astronomical research, saving time and resources.

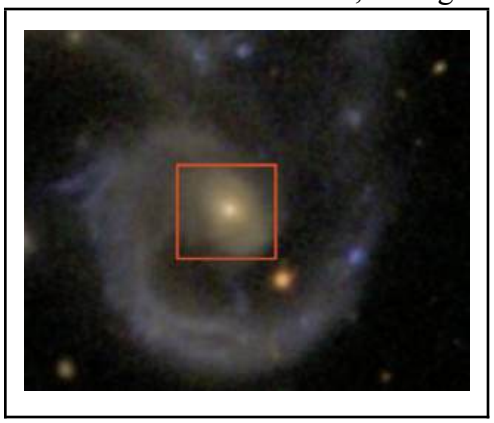


Figure 5: Region of high gas concentration in merger

The main sequence of star-forming galaxies (SFMS) is an observed relationship between stellar mass (Mstar) and star formation rate (SFR) for galaxies that are actively forming stars. It describes how the majority of star-forming galaxies follow a nearly linear correlation between their SFR and stellar mass, meaning more massive galaxies tend to form stars at a higher rate. The SFMS is typically expressed as a power-law relation[6]:

$$\log(SFR) = \alpha \log(M*) + \beta$$

where:

$M*$ is the stellar mass of the galaxy (in solar masses, $M\odot$)

α (slope) is describes how SFR scales with stellar mass (often ~0.7 to 1.0)

B (intercept) depends on the cosmic epoch and specific dataset

The slope was set to 0.7 to suggest that more massive galaxies have slightly lower SFRs

than lower-mass galaxies—meaning they form stars less efficiently relative to their mass. The intercept
of -7 ensures that at M* = 1010 MO, the observed SFR is close to 1 MO/yr. Starburst galaxies lie significantly above the main sequence, so the threshold for starburst formation from the SDSS data was 5 times the SFMS.

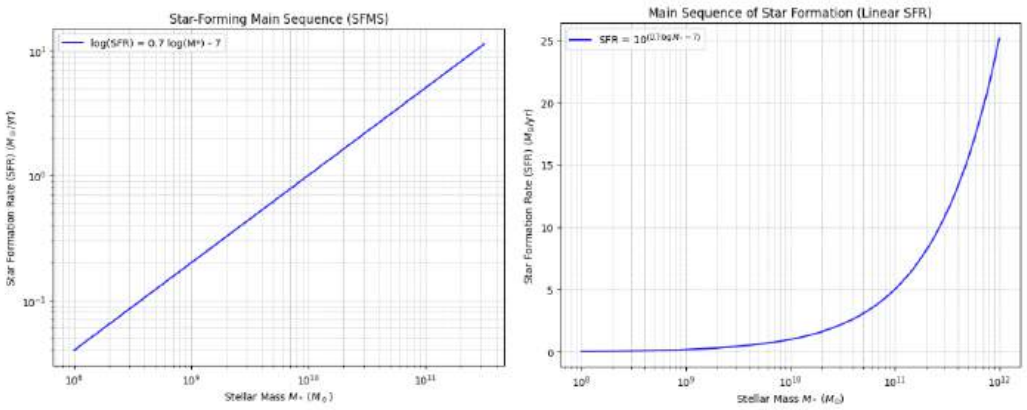


Figure 6: Correlation between SFR and Stellar Mass

Conclusion / Future Research

Since dark matter is theorized to shape the structure of the universe, understanding its role and connection to star formation is vital to the future of astronomy. Developing an efficient and accurate algorithm for classifying starbursts, such as BurstDetector, presents a unique opportunity to improve the accuracy of research in galaxy evolution and the role dark matter halos play in the universe. Starbursts, rapid periods of intense star formation, are often observed in galaxies influenced by the gravitational pull of dark matter halos which can trigger these explosive bursts of star activity. Missions such as the James Webb Space Telescope (JWST), Nancy Grace Roman Space Telescope, and upcoming NASA supported deep-sky surveys aim to study the structure and evolution of galaxies across cosmic time. Integrating our classification model, BurstDetector, with data from these missions could assist astronomers understand how dark matter-driven interactions shaped early galaxy formation. Additionally, our solution can be deployed to analyze data from NASA's Euclid mission, which is dedicated to studying dark energy and dark matter by mapping billions of galaxies.

Overall, our model BurstDetector is a necessity in today's world where classifications of star bursts are manual, slow, and inaccurate. BurstDetector not only opens up new possibilities for scientists to shift their focus and time from classifying starbursts to researching them, but also allows computers of any GPU to perform these classifications. This results in discoveries relating to the field of physics and astronomy being made at a much faster rate.

Works Cited

[1] Pearson, W. J., Wang, L., Trayford, J. W., Petrillo, C. E., & Van Der Tak, F. F. S. (2019). Identifying galaxy mergers in observations and simulations with deep learning. Astronomy & Astrophysics, 626, A49. https://doi.org/10.1051/0004-6361/201935355 Accessed 21 Dec. 2024.

[2] SDSS SkyServer DR12, skyserver.sdss.org/dr12/en/tools/search/sql.aspx. Accessed 3 Jan. 2025.

[3] Shorten, Connor, and Taghi M. Khoshgoftaar. ”A survey on image data augmentation for deep learning.” Journal of big data 6.1 (2019): 1-48. https://link.springer.com/article/10.1186/s40537-019-0197-0?code=b33ae7db-07a1-485c ac3b-4f409f373507 Accessed 5 Jan. 2025

[4] Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 6308-6325. https://doi.org/10.1109/JSTARS.2020.3026724 Accessed 26 Dec. 2024.

[5] Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 5(23), 495. https://cs.nju.edu.cn/wujx/paper/CNN.pdf Accessed 26 Dec. 2024.

[6] Pearson, W. J., Wang, L., Alpaslan, M., Baldry, I., Bilicki, M., Brown, M. J. I., ... & van Der Tak, F. F. S. (2019). Effect of galaxy mergers on star-formation rates. Astronomy & Astro physics, 631, A51. https://doi.org/10.1051/0004-6361/201936337 Accessed 17 Dec. 2024.

Urbanization and Avian Mortality: An Exploratory Analysis of Window-Strike Collisions
By Min Woo Dong

Abstract

Urbanization increases the prevalence of large glass surfaces in cities, which in turn drives high rates of bird-window collisions. This exploratory study uses three complementary datasets—(1) weekly collision counts at a West Coast urban park museum (2008-2014), (2) façade‑level collision surveys in Stillwater, OK (2015-2016), and (3) daily collision totals in Chicago and Cleveland paired with nighttime light intensity—to examine how mitigation, window‑area, and urban density affect avian mortality. A shading retrofit at the Golden Gate Park museum reduced mean weekly strikes from 1.96 to 1.58 ($t = 2.60$, $p = 0.010$) (Kahle et al.). In Stillwater, façades with above-median glass area experienced a mean of 3.72 effort-adjusted bird–window collisions per façade over the two-year survey, versus 0.52 for smaller façades ($t = 3.36$, $p = 0.001$) (Foley). Finally, ambient light level showed a statistically significant but negligible correlation with daily collisions ($r = 0.05$, $p = 0.020$) (Winger et al.). Together, these results highlight the effectiveness of simple mitigation and the outsized role of uninterrupted glass in collision risk, while suggesting that coarse urban‑density proxies need refinement.

Introduction

Bird-window collisions kill up to one billion birds annually in North America alone (Loss et al. 9). As cities grow denser and glass façades proliferate, this problem is expected to worsen. Yet despite extensive documentation of collision hotspots, it is still unclear on which urban features matter most—and how simple interventions can reduce mortality. This paper takes a "micro‑to‑macro" approach, first evaluating the effect of an external shading retrofit at a single museum, then comparing collisions across façades of different glass area, and finally exploring city‑level trends against a proxy for urban density (nighttime light). This paper's goal is to link actionable mitigation with broader urban planning insights.

Literature Review

Bird-window collisions arise primarily because birds cannot distinguish transparent or highly reflective glass from open sky or vegetation (Klem 314). Klem provides a comprehensive review of the sensory and perceptual failures that lead to these strikes, emphasizing that collision risk increases with larger uninterrupted panes and high glass reflectivity (314-15). He further notes that patterned or contrasted surfaces can substantially reduce strikes by increasing window visibility to birds in flight (319-20).

Building on estimates of collision mechanisms, Loss, Will, and Marra quantify the scope of the problem at a continental scale (9). By synthesizing data from multiple monitoring studies, they estimate that up to one billion birds die annually from window strikes in the United States (Loss et al. 9). Their analysis highlights species and habitat vulnerabilities and calls for

standardized reporting protocols to enable robust comparisons of mitigation strategies across regions and building types (12).

Several case studies have applied these insights to real‑world data. The *Bird-Window Collisions at a West-Coast Urban Park Museum* dataset documents fatal strikes over seven years at a single urban park building, offering an ideal before-after framework for assessing retrofit impacts (Kahle et al.). Foley supplements this with a façade‑level survey in Stillwater, OK, where collision counts are linked to measured glass dimensions, allowing investigation of how total window area correlates with mortality (Foley). At a broader scale, Winger et al.'s *Nocturnal Flight-Calling Behaviour Predicts Vulnerability to Artificial Light in Migratory Birds* data set aggregates collision reports from multiple cities and years, which recent studies have paired with remote‑sensing proxies—such as nighttime light intensity—to explore urban‑density effects on strike frequency (Winger et al.).

Despite the wealth of individual studies, few have integrated mitigation experiments, façade design analyses, and macro‑scale density proxies into a unified framework. Most work either tests a single intervention or documents broad mortality patterns, but rarely both. This gap motivates this paper's exploratory "micro‑to‑macro" approach: first evaluating a targeted retrofit, then quantifying the role of window‑area at the façade level, and finally assessing city‑wide trends against an urban‑density proxy. By bridging scales, this paper aims to deliver actionable insights for both building managers and urban planners dedicated to reducing avian mortality.

Data and Methods

This study obtained three publicly available datasets and applied a uniform set of statistical techniques to explore how mitigation measures, façade design, and urban density affect bird-window collisions. Fatal collision records from the Golden Gate Park museum (2008-2014) were obtained from the *Bird-Window Collisions at a West-Coast Urban Park Museum* dataset (Kahle et al.). Façade survey data for Stillwater, Oklahoma (2015-2016) were sourced from Dryad (Foley). City-level collision logs for Chicago and Cleveland were downloaded from the Dryad Digital Repository (Winger et al.) and paired with nightly remote-sensing measurements of light intensity (also Dryad), which served as a proxy for urban density. All data cleaning, aggregation, statistical tests, and figures were produced in Python 3 using pandas for data manipulation, SciPy for inferential statistics, and Matplotlib for visual output.

For the Golden Gate Park analysis, weekly strike counts from January 2008 through December 2014 were aggregated and March 22, 2011—the date of the external shading retrofit—was designated as the cutoff between pre and post retrofit periods. A two-sample Welch's t-test was used to compare mean weekly strikes before and after the retrofit, testing whether the shading treatment produced a statistically significant reduction in collision rates.

In the Stillwater façade survey, each building face's total glass area was calculated by multiplying its length, height, and the proportion of that area covered by glass. Façades were then split into "large" and "small" groups based on the median glass-area value. For each façade,

we took the raw total of observed collisions, divided by that façade's survey effort (number of surveys), and then multiplied by the mean survey count across all façades. This produces an "effort adjusted collision count" representing collisions per unit effort, allowing direct comparison across façades with differing survey frequencies. Within each group, mean and standard deviation of the effort-adjusted collision counts were computed. An independent two-sample Welch's t-test assessed whether the difference in mean collisions between large-area and small-area façades was statistically significant.

At the city scale, daily collision logs from Chicago and Cleveland were concatenated and tallied by date to produce a unified time series of daily strike counts. Nightly light-intensity values for each date were averaged and merged the two series on the date field. Pearson's correlation coefficient was then calculated to quantify the linear association between daily collision counts and average light levels, with significance evaluated via a two-tailed test.

Results

Period	Weeks (n)	Mean Strikes per Week
Pre-retrofit	100	1.96
Post-retrofit	71	1.58

Table 1: Weekly fatal bird-window collision rates at the Golden Gate Park museum before (100 weeks) and after (71 weeks) the 22 March 2011 exterior-shading retrofit, showing a drop in mean strikes from 1.96 to 1.58 per week

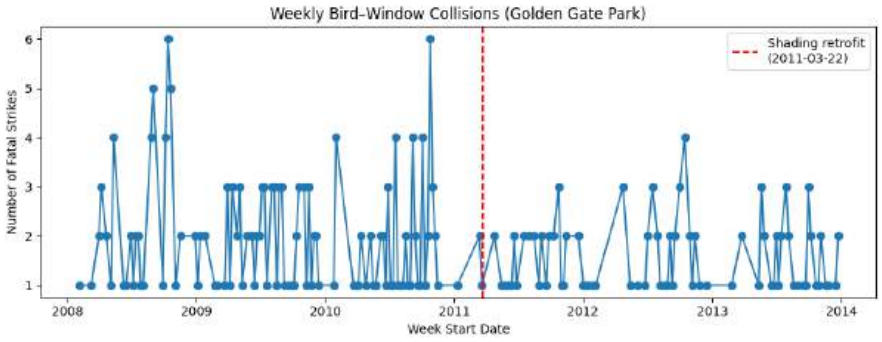


Figure 1: Weekly fatal bird-window collisions at the Golden Gate Park museum, 2008-2014; the dashed vertical line marks the 22 March 2011 shading retrofit that produced a 19% reduction in strikes

Figure 1 plots weekly fatal bird-window collisions at the Golden Gate Park museum from 2008 through 2014, with a vertical dashed line marking the shading retrofit on March 22, 2011. In the 100 weeks before the retrofit, the site averaged 1.96 strikes per week; this fell to 1.58 strikes per week over the 71 weeks that followed (Table 1). A two-sample Welch's t-test confirms that this 19% reduction is statistically significant ($t = 2.60$, $p = 0.010$), indicating that the exterior shading treatment effectively lowered collision rates.

Window-Area Group	N (façades)	Mean Collisions	SD
Large (> median)	12	3.72	1.15
Small (≤ median)	12	0.52	0.34

Table 2: Effort-adjusted collision counts for 24 building façades in Stillwater, Oklahoma, split at the median glass-area value: large-glass façades averaged 3.72 collisions (SD = 1.15) versus 0.52 (SD = 0.34) for small-glass façades

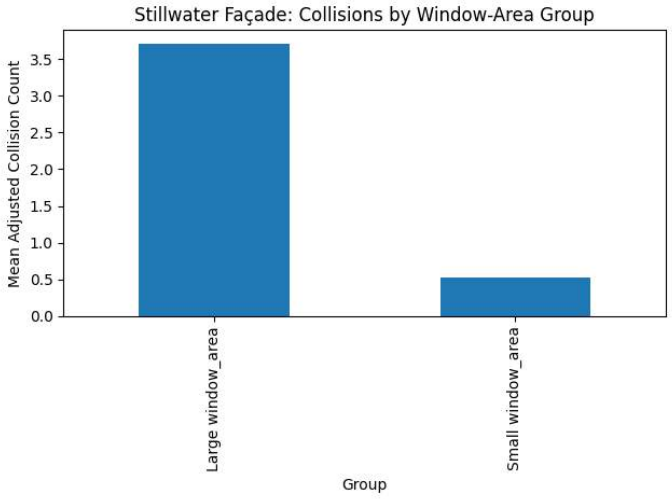


Figure 2: Mean effort-adjusted collision counts for Stillwater façades grouped by total glass area, illustrating the higher risk associated with uninterrupted expanses of glass

Figure 2 shows effort-adjusted collision counts for 24 façades in Stillwater, OK, divided into groups above and below the median glass‑area value. Façades with larger uninterrupted glass expanses experienced an average of 3.72 collisions (SD = 1.15, n = 12), whereas smaller‑glass façades averaged only 0.52 collisions (SD = 0.34, n = 12) (Table 2). The difference between these groups is highly significant (t = 3.36, p = 0.001), underscoring that total window‑area is a critical predictor of avian mortality risk.

Statistic	
Pearson r	0.05
p-value	0.020
Sample size	730

Table 3: Pearson correlation statistics linking daily collision totals (Chicago + Cleveland, n = 730) with nightly average light intensity—a proxy for urban density—yielding r = 0.05, p = 0.020

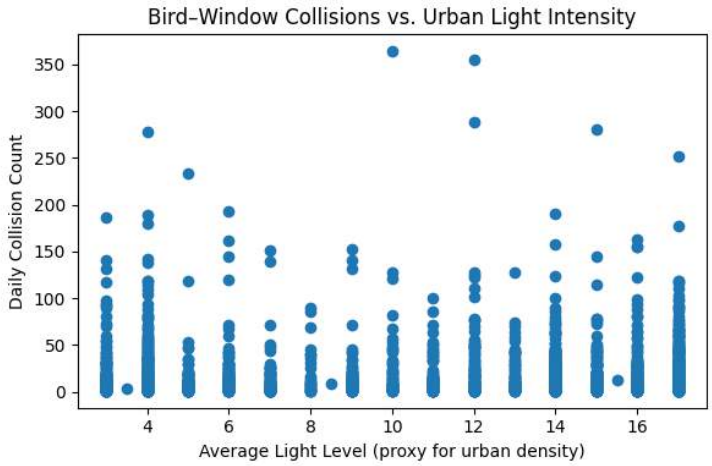


Figure 3: Scatterplot of daily collision counts versus average nighttime light intensity for Chicago and Cleveland combined; the weak positive slope (r = 0.05) indicates that coarse light-based density measures explain little of day-to-day collision variability

Figure 3 presents a scatterplot of daily collision counts against average nighttime light levels—a proxy for urban density—for Chicago and Cleveland combined. Across 730 days, the Pearson correlation coefficient is r = 0.05 with p = 0.020 (Table 3), revealing a statistically significant but practically negligible positive relationship. This weak association suggests that ambient light intensity alone does not capture the full suite of factors—such as migration timing, weather, and building design—that drive day-to-day variation in collision rates.

Taken together, these results demonstrate that both targeted mitigation measures (e.g., shading retrofits) and fundamental façade characteristics (e.g., uninterrupted glass area) have strong, quantifiable impacts on bird-window collisions, whereas coarse city-scale proxies for urban density yield limited explanatory power.

Discussion

The multi-scale analysis performed demonstrates that both targeted interventions and fundamental façade characteristics play crucial roles in reducing bird-window collisions, while coarse proxies for urban density alone explain little of the day-to-day variation. At the smallest scale, the 19% decline in weekly strikes at the Golden Gate Park museum following the March 22, 2011 shading retrofit confirms that simple, low-cost exterior screens can substantially lower avian mortality. This result aligns with Klem's emphasis on contrast and visibility: by interrupting large expanses of reflective glass, the retrofit appears to have made the windows more apparent to birds in flight (Klem 319-20).

At the building scale, the comparison of Stillwater façades underscores the outsized effect of uninterrupted glass area. Façades with above-median glass area experienced, on average, over

seven times more adjusted collisions than those with smaller panes (3.72 vs. 0.52 collisions). This finding suggests that design guidelines should prioritize breaking up large glass surfaces—whether through mullions, fritted patterns, or applied decals—to reduce strike risk. Architects and building managers can apply these insights not only to new construction but also to retrofit projects on existing structures.

By contrast, the city‑level examination of daily collisions versus nighttime light intensity yielded a statistically significant yet small correlation ($r = 0.05$). This weak association indicates that ambient light alone does not capture the complex interplay of factors—such as migration timing, seasonal weather, and species‑specific behaviors—that drive collision rates. It also highlights the limitations of using single-variable proxies for urban density. Future studies could improve on this approach by aggregating collision and light data at coarser time scales (e.g., weekly), incorporating radar-based measurements of migration, or high-resolution satellite images to calculate how much of each area is actually built up

Despite these promising findings, this study has several limitations. The museum dataset represents a single building in a unique microclimate, and the Stillwater façade survey covers only 24 sites, which may limit generalizability. Reporting protocols and data quality likely vary between the Chicago and Cleveland records, potentially introducing noise into the city-level analysis. Finally, this paper's exploratory design does not account for potential confounders such as weather conditions or bird species composition.

To build on this work, three avenues are recommended for further research. First, long-term monitoring of additional retrofit projects across diverse climates would assess the durability of shading treatments. Second, controlled façade-scale experiments—randomly assigning pattern types or densities—could identify optimal configurations for minimizing strikes. Third, integrated urban models that combine collision logs with GIS land-cover, migration-radar, and weather data could yield predictive maps of high-risk zones, guiding both policy and design interventions. By linking micro-scale mitigation to macro-scale urban planning, future efforts can more effectively safeguard migratory and resident bird populations against the growing threat of window collisions.

Works Cited

Foley, Luke E. Stillwater Bird-Window Collision Data. Dryad Digital Repository, 2018. doi:10.5061/dryad.gb5mkkwkf.

Kahle, Logan Q., Maureen E. Flannery, and John P. Dumbacher. Window Strike Data. PLOS ONE, supplement S1, 5 Jan. 2016. doi:10.1371/journal.pone.0144600.s001.

Klem, Daniel, Jr. "Preventing Bird–Window Collisions." The Wilson Journal of Ornithology, vol. 121, no. 2, May 2009, pp. 314–321.

Loss, Scott R., Tom Will, and Peter P. Marra. "Bird-Building Collisions in the United States: Estimates of Annual Mortality and Species Vulnerability." The Condor, vol. 116, no. 1, 2014, pp. 8-23.

Winger, Benjamin M., Brian C. Weeks, Andrew Farnsworth, Andrew W. Jones, Mary Hennen, and David E. Willard. Nocturnal Flight-Calling Behaviour Predicts Vulnerability to Artificial Light in Migratory Birds. Dryad Digital Repository, 11 Mar. 2019, doi.org/10.5061/dryad.8rr0498.

Variation of Vaccine Effectiveness Across Different Age Groups By Ishan Wadhwani

Background

Vaccination is a safe, simple, accessible, and effective method to protect against harmful diseases before the disease enters the body ("Vaccines and Immunization: What Is Vaccination?"). It harnesses the human body's natural defenses to create resistance to infections by introducing an antigen into the human body, emulating an infection, and triggers the immune system to respond (World Health Organization, "Vaccine Efficacy, Effectiveness and Protection"). Vaccines are effective against numerous diseases, including the common Flu, Dengue, Polio, and, more recently, COVID-19. Vaccination is extremely accessible as well, with billions of people vaccinated worldwide against a multitude of diseases. For instance, more than 5.55 billion people have been vaccinated with one of the COVID-19 vaccines, reiterating its importance as one of the most important medicinal products (Bengali). Recognizing the invading germ (eg. virus), it produces antibodies, naturally produced proteins to fight disease and remembers the specific disease and how to counter it, minimizing the effects of future exposure to the disease (Kurzgesagt). This immunization is achieved through the complex interconnected efforts of various Leukocytes, white blood cells that are part of the immune system that include B cells and T cells ("Immune Cells"). Vaccine effectiveness, a measure of how well vaccines work, is a crucial aspect of the immunization process (World Health Organization, "Vaccine Efficacy, Effectiveness and Protection"). The effectiveness of a vaccine includes the immune response triggered by the introduction of the vaccine induced antigen, as well as further immunological reactions that support the core immune response (Chen et al.). While vaccination as a whole does seem extremely effective, the effectiveness of vaccination varies depending on numerous factors. The age of the patient, for instance, contributes significantly to the effectiveness of a vaccine as many factors such as the maturity of the immune system and the gradual decline in immune functionality due to aging, prominently known as immunosenescence, influence vaccination effectiveness (Lord). This yields the following research focus: To what extent does vaccine effectiveness vary across different age groups?

Vaccine Production: Enzymes

To answer this question, it is important to understand the production of vaccines and the role of enzymes in this process. One type of vaccination that we can analyze are mRNA vaccines. Generally, mRNA vaccines are produced in the following series of steps: 1) DNA templates are reproduced based on the template strand, often through DNA replication, 2) RNA Polymerase and other enzymes transform the double-stranded DNA to single stranded mRNA in transcription, 3) Synthesized mRNA undergoes comprehensive quality control protocols, and 4) mRNA is encapsulated in lipid nanoparticles to protect the mRNA (Chaudhary et al.). Specific mRNA vaccines are produced with various added steps, such as common COVID-19 mRNA vaccines, which additionally involves a spike protein to create a specific antigen that primes the immune system to create antibodies against this specific antigen (Chaudhary et al.). Regardless

of the type of vaccine, enzymes play a critical role in vaccine manufacturing, in the transcription, modification, and quality assurance processes (Fang et al.). In transcription, enzymes transcribe the genetic information from the 3’ to 5’ DNA template strand and synthesizes RNA in the 5’ to 3’ direction. Post-transcriptional modifications, such as DNA splicing, 3’ end polyadenylation, and 5’ end cap addition, are conducted with Transposases, polymerase, and other enzymes to precisely orchestrate the modifications (Clara). Lastly, raw enzymes eliminate unwanted RNA fragments and/or impurities in mRNA quality control (Clara). The aforementioned processes that enzymes are involved in are critical to the accurate, efficient, and safety of mRNA content in mRNA vaccines.

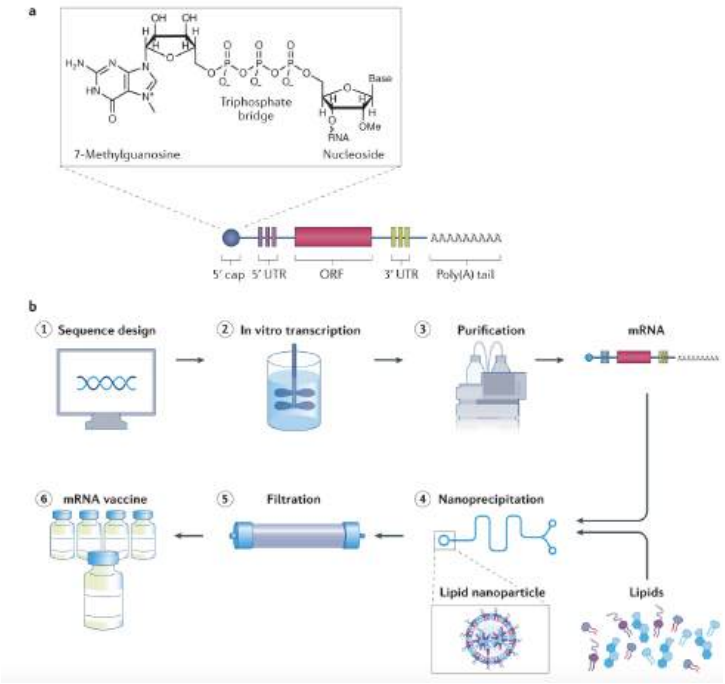


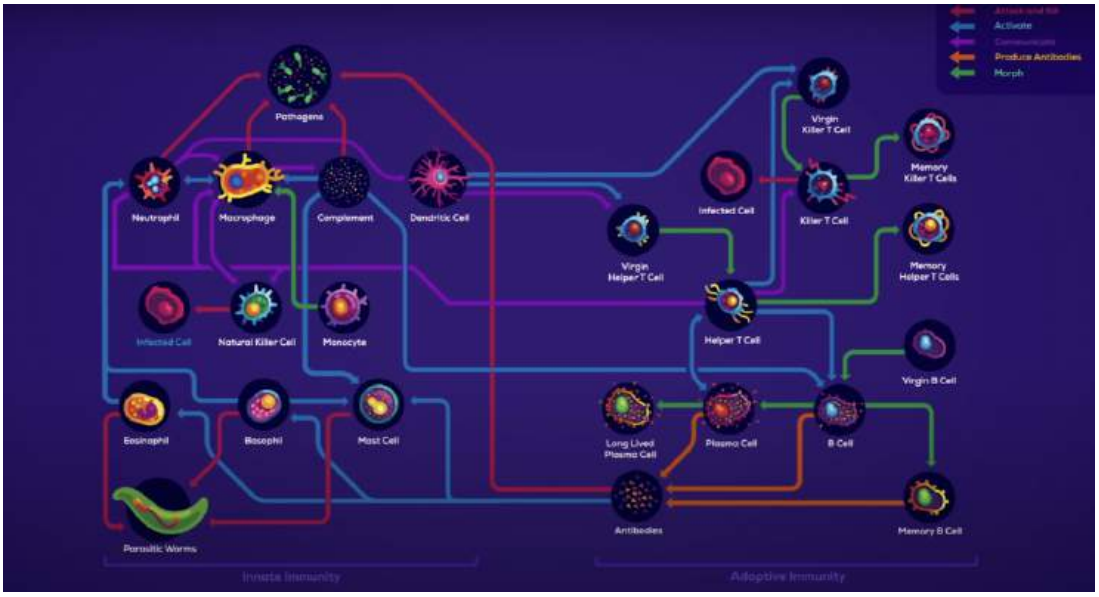
Fig 1: In vitro-transcribed (IVT) mRNA contains five structural elements: a 5′ cap, 5′ and 3′ untranslated regions, an open reading frame, and a poly(A) tail, is synthesized from plasmid DNA, undergoes purification, is encapsulated in liquid nanoparticles and filtered, and stored in sterilized vials for vaccination

As we age, a multitude of changes occur to enzymes, albeit to different extents depending on the enzyme (Cook; Debès et al.). This caveat occurs due to distinct chemical characteristics and their designated purpose, especially in terms of vaccination production. For RNA polymerase, a major enzyme involved in mRNA vaccine production, transcriptional elongation speed typically increases with age; that is, the speed at which RNA polymerase can synthesize and elongate an mRNA strand (Cook). However, an increase in age (and therefore speed) also leads to a decrease in genomic accuracy (Debès et al.). These enzymatic changes have significant repercussions on overall vaccine effectiveness. As RNA polymerase becomes less accurate, various changes transcending transcription in vaccine production occur as a result: the mRNA strand has a greater chance of genetic mutation due to increased inaccuracies (transcription);

enzymatic conducted genomic modifications will not be completed as intended (modification); and RNA and DNA impurities will be ineffectively eradicated from the final mRNA vaccine content (Cook; Debès et al.). These age-induced changes at a broader level will lead to a less accurate and perhaps even incorrect vaccine as a whole, making the vaccine ineffective as a whole.

Fig 2: A broad overview of the Immune system, characterizing cells based on adaptive or innate immunity

Immunosenescence Implications on Vaccine Effectiveness: Biochemistry and Cell Biology



As vaccines work by stimulating the immune system without causing the disease, the immune system of the vaccinated person is exposed to the antigen and launches the immune response to effectively tackle the foreign substance(s) (World Health Organization, "Vaccine Efficacy, Effectiveness and Protection"). Numerous immune cells and organs are part of this immune response, which are separated into two major sections in the immune response: Innate immunity, cells involved in the "first line of defense," and adaptive immunity, cells involved in the destruction of pathogens based on alerts from innate immune cells ("Immune Cells"; Kurzgesagt). Innate immune cells include: Macrophages, which are the initial immune response cells that to degrade and ingest bacteria and alert other immune cells of the antigen; and Neutrophils, which launch a combination of lethal chemicals including NADPH oxidase-derived reactive oxygen species (ROS), cytotoxic granule components, antimicrobial peptides, and neutrophil extracellular traps (NET) at bacteria, damaging and killing them (Teng et al.; Parker et al.). After this innate immune response, adaptive immune cells begin working; T cells (produced in Thymus and subdivided into helper T cells (CD4+), cytotoxic T cells (CD8+), memory helper T cells, etc.), which are a type of lymphocyte that, on a broad level, identify and destroy bacteria and pathogens; B cells (produced in bone marrow), which have a central role in humoral immunity and producing vaccines and presenting them to T-Cells for further immunization (NIH;

Kurzgesagt). While general immune frameworks can be followed, it is important to note that the appropriate immune response and the cells that are part of it also depend on the type of vaccine and the antigen introduced.

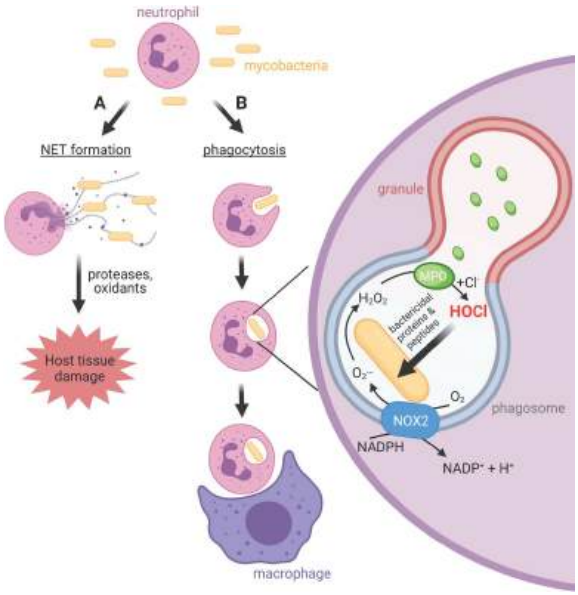


Fig 3: Neutrophil's release of NETs, neutrophil extracellular traps, combats mycobacteria, phagocytizing these bacteria, and transferring live bacteria or antimicrobial agents to macrophages

Immunosenescence, the gradual decline in immune function with age, is primarily driven by reduced biochemical capabilities as well as weaker immune cells (Liu et al.). Lymph nodes, which filter out harmful substances and stimulate immune responses after vaccination, are significantly altered following reduced germinal centers (in lymphoid tissue) and a decrease in the number of follicular dendritic cells and fibroblastic reticular cells; cells instrumental in the innate immune response (Lord; Liu et al.). Aging also facilitates reduced B cell production from the bone marrow cause, flawed affinity maturation, thereby weakening the adaptive immune response to vaccine-induced antigens (Chen et al.). Furthermore, T cell proliferation and activation is extremely reduced with age (for most T cells) following Thymic Involution, the shrinking of the thymus due to age-related reasons (Lord). This, in turn, contributes to decreased helper T cells that provide imperative differentiation signals to B cells in adaptive immunity. Moreover, neutrophils undergo impaired ROS production and phagocytosis, reducing their affinity to ingest bacteria (Ogawa et al.; Parker et al.).

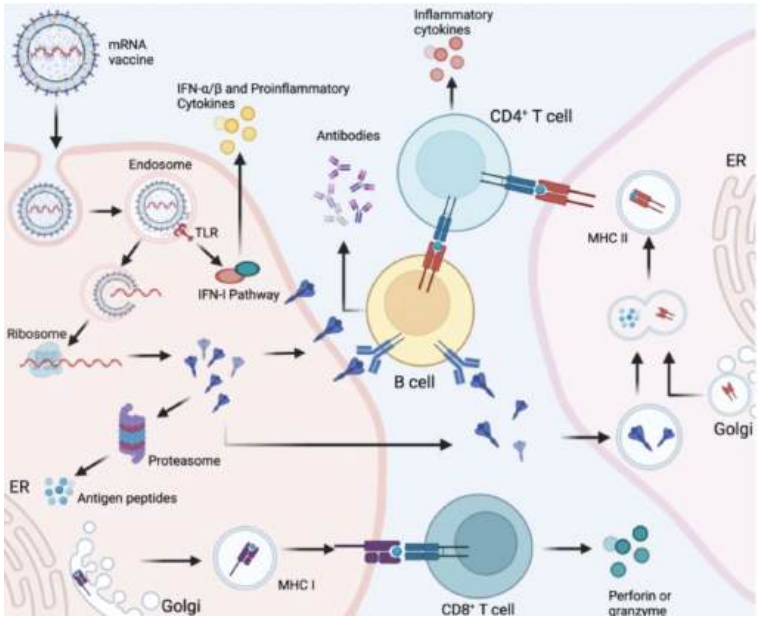


Fig 4: mRNA delivery through an mRNA vaccination through endocytosis and translation into protein by ribosomes

Despite memory T cells and B cells increasing in numbers with age, their function in the immune response to antigens significantly decreases (Liu et al.; Lord). These multifaceted cellular and chemical changes contribute to immunosenescence and lead to the degradation of immune functionality over time. In terms of vaccination, this means that the intended immune response will be less effective to counter the antigen introduced by the vaccine, leading to limited protection from the disease and/or unintended side effects for the patient.

Case Study: COVID19

COVID-19, a disease caused by the virus SARS-CoV-2, is a contagious disease that was discovered in late 2019 and classified as a pandemic between 2020-2023 (World Health Organization, Coronavirus). The Pfizer-BioNTech COVID-19 Vaccine, along with various other vaccines aiming to tackle the pandemic, including the Moderna COVID-19 Vaccine, were developed later that year and early 2021. While these vaccines were overall quite effective in countering the disease, there are some notable relationships between age and the effectiveness of the vaccine. A research study, “Age and product dependent vaccine effectiveness against SARS-CoV-2 infection and hospitalisation among adults in Norway: a national cohort study, July–November 2021” conducted by Jostein Starrfelt and associates explored the impact of the dose number on vaccine effectiveness and chance of hospitalization between different age groups (Starrfelt et al., “Age and Product Dependent…”). Their study concluded that vaccine effectiveness against COVID-19 decreased by a cumulative average across age groups from 81.3% after 2-9 weeks after the second dose to 8.6% after 33 weeks respectively. More specifically, vaccine effectiveness was highest among 18-44 year olds after 2-9 weeks at 83.2%

(with confidence interval 82.6 to 83.8) compared to those between 45-64 years old at 75.6% (with confidence interval: 74.1 to 77.0). After 33 weeks after the second dose, vaccine protection was extremely limited with 5.2% for 18-44 year olds (confidence interval: -1.9-11.8) compared to 0.5% for 44-64 year olds (confidence interval: -9.4-9.5) (Starrfelt et al., "Fig. 2…").

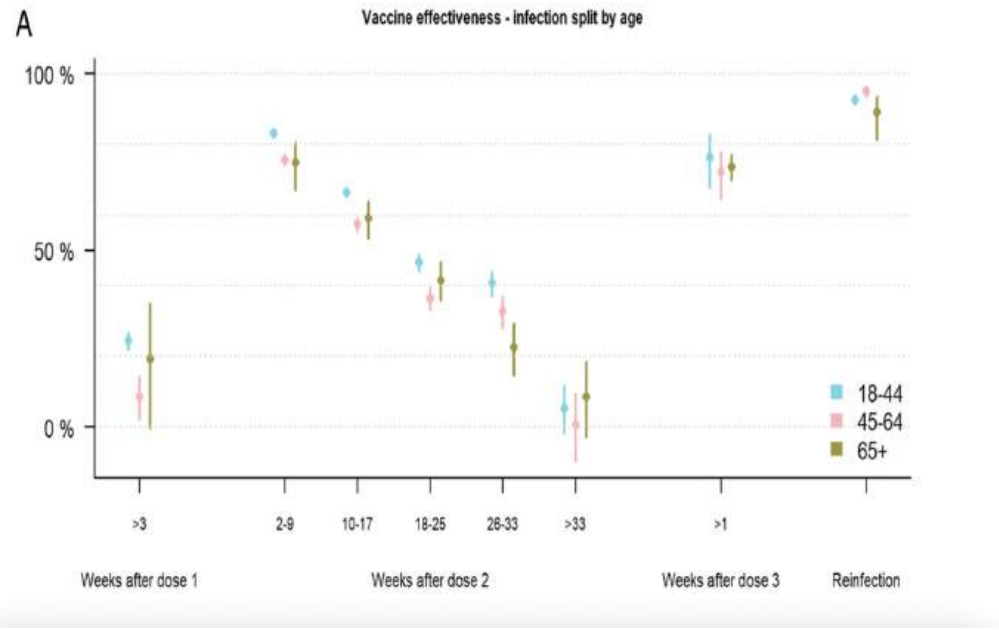


Fig 5: Norwegian Results Chart: Vaccine effectiveness shown against infection for age 18–44 years (blue), 45 to 64 (pink) and 65+ years (green) among Norwegian adults (July-November 2021). Adjusted for a variety of factors (age, sex, living conditions, etc.)

Another research study, "The Impact of Age Difference on the Efficacy and Safety of COVID-19 Vaccines: A Systematic Review and Meta-Analysis," conducted by Jiting Wang from The Affiliated Hospital of Southwest Medical University, Luzhou China, and associates compared the efficacy and safety of five different COVID-19 vaccines depending on the age group (Wang et al.). Their meta-analysis yielded results that the virus geometric mean titer (GMT) values [a higher GMT value is associated with increased immunogenicity and vice versa] were significantly higher among younger individuals compared to older individuals. However, the "safety" of younger individuals was lower, meaning they were more prone to various AEs (Adverse Effects) including local muscle pain, rashes, and fever, which were tolerable among the subset. The study attributed immunosenescence as the primary reason for the GMT values of older individuals to be lower. Depending on the vaccine, different effectiveness rates for different age groups can be identified. For example, the Pfizer-Biotech vaccine yielded a 95% effectiveness for people 16 and older and 94% for people age 65 and older (Pfizer and BioNTech Conclude Phase 3 Study of COVID-19 Vaccine Candidate, Meeting All Primary Efficacy Endpoints | Pfizer), whereas the Moderna vaccine yielded 95.6% effectiveness for people ages 18 to 65 and 86.4% for people older than 65 (Moderna COVID-19 Vaccine VRBPAC Briefing

Document). Despite the type of vaccine and even special effects for COVID-19 (eg. inflamed respiratory tract in lungs, exacerbated by older individuals), the overall trend revealed by the data aligns with scientific evidence on immunosenescence and enzymatic, cellular, and biochemical changes occurring as a result of age increase (Lord; Chen et al.).

Conclusion

In conclusion, an overall increase in age has tended to correlate to a significant decrease in overall vaccine effectiveness due to changes in enzymatic activity, biochemical and cellular responses, and immunosenescence. Enzymes such as RNA polymerase likely have reduced accuracy with aging, despite faster transcriptional elongation speeds. Biochemical processes, including ROS production of neutrophils, are limited as we age, reducing the ability of immune cells to ingest and degrade bacteria and therefore hindering the immune response and the intended effect of a vaccine. Immune cells undergo various immunological changes in their functions, such as T cells which face a reduction in their production and activation numbers following the involution of the thymus, which restricts the ability of the antigen-fighting-cell to carry out its specified functions including remembering the disease, leading to an ineffective vaccine and possible unintended side effects for countering the antigen introduced in the vaccine as the immune system is unable to remember how to tackle the antigen upon pathogenic introduction. These biological changes are demonstrated in clinical trials relating to the COVID-19 vaccines, which yield results indicating weaker immunogenicity among older individuals, attributed to immunosenescence. It can therefore be assumed that aging does correlate to a significant reduction in the effectiveness of current vaccines (Wang et al.; Starrfelt et al., "Age and Product Dependent…"). However, it is important to note that this paper focuses on COVID-19 as a case-study, and it is also important to consider other diseases and age ranges to substantiate this claim.

Nevertheless, the world of vaccination is continuously evolving. Pioneered by scientists in the 1990s, DNA vaccination has slowly grown in popularity over time as researchers and immunologists alike seek innovative vaccine solutions that are not affected by immunosenescence (Shukla). DNA vaccines use circular molecules of DNA (plasmids) to introduce a gene that stimulates the immune system. For instance, ZyCoV-D, a 2021 COVID-19 vaccine, includes a gene that codes for the spike protein of SARS-CoV-2. The plasmid enters the human cell and then travels into the nucleus of the patient to undergo transcription into mRNA. The mRNA is then converted to a bacterial/viral protein by various enzymes in cytoplasmic translation, and the immune system elicits the immune response. While this new biotechnology does seem innovative, it is important to consider the effect of aging on DNA. A causal relationship between aging and progressive atrophy has been clinically revealed in that DNA breaks, suffers mutations, and cross-links as a result of aging and degeneration of cells (Maynard et al.). Limited research suggests that DNA breakages and mutations will cause unintended effects on the DNA content in DNA vaccines, leading to alterations in its transcription into mRNA and subsequent translation into a distinct protein. This leads to an immune response

targeted at a different viral antigen than intended. However, further research is needed to scientifically prove this and identify the extent to which immunosenescence impacts the effectiveness of DNA vaccination.

The development of further DNA vaccines is therefore heavily encouraged to understand the complexity of this emerging technology, as well as the extent to which aging affects its function.

Works Cited

Bengali, S. (2022, January 28). The world surpasses 10 billion CoVid vaccine doses administered. The New York Times. https://www.nytimes.com/2022/01/28/world/covid-vaccines-10-billion-doses.html.

Chaudhary, Namit, et al. "IVT mRNA Is Formulated into Lipid Nanoparticle Vaccines Using a Cell-Free Production Pipeline." Nature Reviews Drug Discovery, 21 Sept. 2021, nature.com/articles/s41573-021-00283-5. Accessed 11 Aug. 2025.

Chen, Judy, et al. "How Aging Impacts Vaccine Efficacy: Known Molecular and Cellular Mechanisms and Future Directions." Trends in Molecular Medicine, vol. 28, no. 12, 7 Oct. 2022, pp. 1100–1111, doi:10.1016/j.molmed.2022.09.008.

Clara, Lisa. "The Significance of Raw Enzymes in the Success of mRNA Vaccines." PharmiWeb, 22 Sept. 2023, pharmiweb.com/article/the-significance-of-raw-enzymes-in-the-success-of-mrna-vaccines. Accessed 11 Aug. 2025.

Cook, Jane. "Unlocking the Secrets of Cellular Aging: The Role of RNA Polymerase II." Bridge Informatics, 13 Apr. 2023, bridgeinformatics.com/unlocking-the-secrets-of-cellular-aging-the-role-of-rna-polymerase-ii/. Accessed 11 Aug. 2025.

Debès, Cédric, et al. "Ageing-Associated Changes in Transcriptional Elongation Influence Longevity." Nature, vol. 616, no. 7958, 2023, pp. 814–821, doi:10.1038/s41586-023-05922-y.

Fang, E., et al. "Cellular and Humoral Immune Responses Induced by Messenger RNA (mRNA) Vaccine." Signal Transduction and Targeted Therapy, 2022, doi:10.1038/s41392-022-00950-y. Accessed 11 Aug. 2025.

"Immune Cells." National Institutes of Health, 14 Jan. 2014, nih.gov/news-events/nih-research-matters/immune-cells. Accessed 11 Aug. 2025.

Kurzgesagt. How the Immune System ACTUALLY Works – IMMUNE. YouTube, 10 Aug. 2021, youtube.com/watch?v=lXfEK8G8CUI. Accessed 11 Aug. 2025.

Liu, Z., et al. "Immunosenescence: Molecular Mechanisms and Diseases." Signal Transduction and Targeted Therapy, vol. 8, 200, 2023, doi:10.1038/s41392-023-01451-2.

Lord, Janet M. "The Effect of Ageing of the Immune System on Vaccination Responses." Human Vaccines & Immunotherapeutics, vol. 9, no. 6, 2013, pp. 1364–1367, doi:10.4161/hv.24696.

Maynard, Scott, et al. "DNA Damage, DNA Repair, Aging, and Neurodegeneration." Cold Spring Harbor Perspectives in Medicine, vol. 5, no. 10, 2015, doi:10.1101/cshperspect.a025130. Accessed 11 Aug. 2025.

Moderna COVID-19 Vaccine VRBPAC Briefing Document. 2020, www.fda.gov/files/advisory%20committees/published/VRBPAC-12.17.20-Meeting-Briefing-Document-FDA_0.pdf.

Ogawa, K., et al. “The Association of Elevated Reactive Oxygen Species Levels from Neutrophils with Low-Grade Inflammation in the Elderly.” Immunity & Ageing, vol. 5, no. 13, 2008, doi:10.1186/1742-4933-5-13.

Parker, Heather A., et al. “Neutrophil Activities at Sites of Mycobacterial Infections.” Frontiers in Immunology, 23 Dec. 2021, frontiersin.org/articles/10.3389/fimmu.2021.782495/full. Accessed 11 Aug. 2025.

Pfizer and BioNTech Conclude Phase 3 Study of COVID-19 Vaccine Candidate, Meeting All Primary Efficacy Endpoints | Pfizer. www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-conclude-phase-3-study-covid-19-vaccine.

Shukla, Deep. “DNA vs. mRNA Vaccines: Similarities and Differences.” Medical News Today, 28 Sept. 2021, medicalnewstoday.com/articles/dna-vs-mrna-vaccines-similarities-and-differences. Accessed 11 Aug. 2025.

Starrfelt, J., et al. “Age and Product Dependent Vaccine Effectiveness Against SARS-CoV-2 Infection and Hospitalisation Among Adults in Norway: A National Cohort Study, July–November 2021.” BMC Medicine, vol. 20, 278, 2022, doi:10.1186/s12916-022-02480-4.

Starrfelt, J., et al. “Fig. 2, July–November 2021.” BMC Medicine, vol. 20, 278, 2022, doi:10.1186/s12916-022-02480-4.

Teng, Tie-Shan, et al. “Neutrophils and Immunity: From Bactericidal Action to Being Conquered.” Journal of Immunology Research, 2017, doi:10.1155/2017/9671604.

Vaccine Efficacy, Effectiveness and Protection. World Health Organization, 14 July 2021, who.int/news-room/feature-stories/detail/vaccine-efficacy-effectiveness-and-protection.

“Vaccines and Immunization: What Is Vaccination?” World Health Organization, 30 Aug. 2021, www.who.int/news-room/questions-and-answers/item/vaccines-and-immunization-what-is-vaccination.

Wang, Jiting, et al. “The Impact of Age Difference on the Efficacy and Safety of COVID-19 Vaccines: A Systematic Review and Meta-Analysis.” Frontiers in Immunology, vol. 12, 2021, 758294, doi:10.3389/fimmu.2021.758294. Accessed 11 Aug. 2025.

From Overwork to Flexibility: Exploring the Potential of Introducing A Four-Day Work Week in Taiwan By Vanessa Tso

Abstract

The study explores the perceptions of young Taiwanese workers (aged 20-35) regarding the potential implementation of a four-day workweek in Taiwan. As global debates around work-life balance and labor reform gain momentum, this research aims to understand whether such a model is viewed as both feasible and beneficial within Taiwan's unique cultural and economic context.

The study employed a quantitative, online survey distributed across Taiwan, collecting responses from 59 participants through convenience sampling. The questionnaire included both closed- and open-ended questions covering demographics, awareness, perceived feasibility, structural preferences, and perceived government support. Data were analyzed using descriptive statistics and chi-square tests of independence to identify significant relationships between key variables.

Findings indicate that educational background significantly influences initial reaction, with participants holding a doctoral degree showing a higher level of skepticism towards the four-day workweek. In contrast, participants with prior experience of flexible or reduced hours reported more positive views on feasibility. Positive first impressions also strongly influenced beliefs about feasibility. Moreover, trust in government support was linked to more optimistic expectations, such as enhanced creativity. Notably, even those who doubted feasibility expressed support for progressive models such as the 32-hour workweek.

These insights suggest that a successful introduction of a four-day workweek in Taiwan would require gradual implementation, strong government leadership, and clear communication to address public doubts and sector-specific concerns.

1. Introduction

The traditional five-day workweek, once considered a standard of modern economies, is increasingly being questioned as global conversations around work-life balance, mental health, and productivity gain attention. Trials of the four-day workweek in countries like Iceland, Japan, New Zealand, and the UK highlight its significant impact. These experiments have shown positive outcomes, such as improved employee well-being, reduced burnout, and sustained or even increased productivity. In Iceland, productivity rose by 1.5%, while UK employees reported greater job satisfaction and fewer sick days (Kobie, 2024; Gecsoyler, 2024). However, critics argue that shorter workweeks may lead to increased workloads, especially in sectors requiring constant staffing, and raise concerns about possible disruptions to service quality or economic efficiency (Vega, 2022). By addressing these pressing concerns, the four-day workweek aligns with the United Nations' Sustainable Development Goals (SDGs), particularly SDG 8: Decent Work and Economic Growth and SDG 3: Good Health and Well-being. As countries work

together to create healthier and more productive work environments, the four-day model emerges as a compelling solution for the challenges of modern labor practices.

Taiwan, a country known for its highly productive yet demanding work culture, faces unique challenges in this context. Driven by global competitiveness, many workers in Taiwan endure long hours and high levels of stress, with three-quarters of Taiwanese employees working long hours, averaging 2,000 hours per year, significantly higher than the OECD average of 1,716 hours (Taipei Times, 2015; HR Asia, 2023). While younger generations in Taiwan are increasingly prioritizing well-being over traditional workplace norms, this shift contrasts with the country's historical emphasis on hard work and long hours as a cultural value (NNRoad, 2023). Globally, there is a growing trend toward flexible work arrangements and shorter workweeks, with many countries embracing these models to enhance productivity and work-life balance. Although the government has introduced labor reforms, such as capping weekly working hours at 40 and enhancing overtime protections, work-life balance remains elusive for many (Ministry of Labor, Taiwan, 2023). Rising stress levels and burnout among workers underscore the urgency of addressing these systemic issues. In this regard, the four-day workweek offers a potential pathway to reshape Taiwan’s labor landscape, balancing productivity with improved quality of life for employees (Wong et al., 2019).

This study investigates the potential of a four-day workweek in Taiwan, focusing on global insights and local perceptions. First, the study will draw on global case studies on experiments of shorter workweeks to evaluate evidence-based insights that might help policymakers, researchers, and organizations in Taiwan anticipate potential outcomes. There is significant value in understanding best practices from successful examples, while also examining the challenges faced by other countries, to help Taiwan prevent similar issues from happening and design proper solutions. Second, this study will explore perceptions among young working individuals aged 20–35, a demographic often grappling with work-life balance challenges while representing a critical segment of Taiwan’s workforce (CIVICUS Monitor, 2023). By examining their views, this research aims to identify the feasibility and potential impact of a four-day workweek in addressing Taiwan's unique labor challenges.

For the purposes of this study, a “four-day workweek” is defined as a work arrangement in which employees fulfill their standard weekly working hours, typically 40 hours, without any reduction in salary, benefits, or entitlements compared to a conventional five-day schedule. This model assumes no loss in total working time, but a redistribution of hours across fewer days. Unless otherwise specified, the four-day arrangement discussed in this research is considered voluntary and non-mandated, meaning it is implemented at the discretion of employers or by mutual agreement between employer and employee, rather than through government legislation. In instances where international models diverge from this framework, such as through compressed hours, pay reductions, sector-specific trials, or mandated state policies, those distinctions are noted accordingly to preserve contextual accuracy and facilitate meaningful cross-national comparison.

2. Literature review

2.1 Global experiences of shortened workweeks

The four-day workweek has been trialled in various countries worldwide; most of them have provided valuable insights into its potential benefits and challenges. Countries like Iceland, the UK, Japan, and New Zealand have conducted significant trials to evaluate shortened work weeks' economic and social impacts. The varied international experiences underscore the necessity for a culturally and economically tailored approach to implementing shorter work-weeks, informing the specific research questions addressed in this study regarding Taiwan's feasibility.

In Iceland, one of the first countries to trial a four-day workweek on a large scale, the results have been overwhelmingly positive. Between 2015 and 2019, trials conducted mainly in the public sector involved a reduction in work hours from 40 to 35 hours per week with no pay cuts (Kobie, 2024). The outcomes showed that productivity increased by 1.5%, and the country's GDP growth outpaced other Nordic nations during the same period. Employees reported significantly reduced stress levels, with 52% noting that it became easier to balance work and personal life (Kobie, 2024). By 2022, 59% of Iceland's workforce had transitioned to shorter working hours, supported by legislative changes. However, the trials also highlighted some concerns about implementation costs, particularly in the public sector, which experienced an estimated 0.11% increase in its budget to accommodate the changes (Kobie, 2024).

In the UK, a six-month trial involving 61 organizations and 2,900 employees provided equally compelling results. The companies that were involved had stable revenues, with an average 1.4% increase, while employee retention improved significantly, with 57% fewer resignations during the trial period (Gecsoyler, 2024). Social benefits were evidenced: 71% of employees reported lower burnout, while sick days decreased by 65%. Additionally, employees found it easier to manage family and social commitments, improving their overall job satisfaction. However, the trial also highlighted sectoral challenges, particularly in industries requiring constant staffing, where compressed workweeks could lead to increased workloads for remaining staff.

In New Zealand and Australia, experiments like the four-day week trial by Perpetual Guardian and Unilever have showcased the potential for shorter workweeks to be successfully implemented in multinational organizations and local businesses. In New Zealand, Perpetual Guardian, a financial services company, piloted a reduced workweek in 2018, allowing employees to work four days while maintaining full pay. The results were overwhelmingly positive, with productivity levels remaining constant while employee engagement, job satisfaction, and work-life balance significantly improved (Hytner, 2022). Inspired by such success, Unilever conducted its trial across regions, demonstrating the scalability of the model in a multinational organization. Employees reported enhanced focus and reduced stress, with many emphasizing the positive impact on their work-life balance (Hytner, 2022). However, concerns

remain about whether productivity gains can be sustained in high-demand industries, where workloads may challenge the feasibility of maintaining such benefits long-term (Hytner, 2022).

In Japan, the four-day workweek has received backing from the government as part of efforts to improve work-life balance and address the nation's declining birth rate. As part of its Work Life Choice Challenge (WLCC) summer activity in August 2019, Microsoft Japan conducted a four-day workweek experiment, allowing employees to work four days while maintaining full pay. The results showed a 40% increase in productivity, as well as reduced electricity usage and office resource costs, highlighting the potential benefits of shorter workweeks (Hurst, 2023). However, deep-seated cultural norms around overwork created significant barriers to widespread adoption. Many employees fear that working fewer hours may be perceived as a lack of dedication, creating hesitancy among workers and employers alike (Hurst, 2023).

In Spain, the government launched a pilot program for a four-day workweek to assess its feasibility and impact on productivity and employee well-being. Funded by a €50 million government initiative, the trial aimed to help small- and medium-sized enterprises (SMEs) implement reduced working hours without cutting employee salaries (Kassam, 2021). Participating companies received financial assistance to mitigate potential risks during the transition. Early reports from companies like Delsol, a software firm that introduced a four-day workweek, showed stable productivity levels and increased employee satisfaction (Vega, 2022). Employees reported better work-life balance and reduced stress, while the company maintained efficiency without significant revenue loss. However, challenges emerged for sectors with high customer demands or limited flexibility, where managing workloads with reduced hours proved difficult (Vega, 2022). Spain's experiment reflects the potential of shorter workweeks but underscores the importance of adopting implementation strategies for different industries (Kassam, 2021).

In Belgium, the government introduced a flexible four-day workweek option as part of the 2022 labor deal aimed at enhancing work-life balance. Unlike traditional four-day workweek models that reduce overall working hours, Belgium's approach allows employees to condense their standard 38-hour workweek into four days instead of five (Van Olmen, 2022). This initiative has been well-received by workers seeking greater control over their schedules, as it offers flexibility without sacrificing income. However, critics argue that this model may not effectively address issues of overwork and burnout since the total hours worked remain unchanged (Van Olmen, 2022). While it represents a step forward in promoting flexibility, the model highlights the ongoing challenge of balancing productivity and employee well-being without reducing workload demands (Van Olmen, 2022).

These global trials show that the four-day workweek can bring significant economic and social benefits, including increased productivity, better employee well-being, and improved work-life balance. However, they also underscore the importance of adjusting and adopting implementation strategies to address unique cultural and economic challenges. These experiences provide valuable insights for understanding how a four-day workweek might be

adapted to Taiwan's specific context, where long hours and high productivity expectations dominate the workplace (HR Asia, 2023). For example, like Japan, Taiwan shares cultural norms that emphasize long working hours as a sign of dedication, which could create resistance among workers and employers (NNRoad, 2023). Additionally, Taiwan's manufacturing and technology sectors, much like high-demand industries in the UK and Belgium, would need tailored approaches to manage workloads while maintaining operational efficiency (Gecsoyler, 2024). Drawing from these international examples, Taiwan could explore gradual implementation, starting with flexible scheduling or sector-specific trials to address these challenges effectively (Ministry of Labor, Taiwan, 2023).

Country	Observed Benefits	Observed Challenges
Iceland	Increased productivity (+1.5%), reduced stress, improved work-life balance, 59% workforce adopted changes	Public sector faced additional costs (0.11% budget increase); scalability concerns
UK	Stable productivity (+1.4%), reduced burnout (71%), fewer sick days (65%), better employee retention (57%)	Sectoral challenges, especially for industries requiring constant staffing
New Zealand	Enhanced focus and reduced stress, improved work-life balance, stable productivity	Questions about sustaining long-term productivity in high-demand industries
Australia	Similar results as New Zealand, showcasing potential scalability across multinational organizations	Limited data on long-term impacts; sector-specific adaptation challenges
Japan	Productivity increased (+40% in Microsoft Japan's trial), reduced electricity usage and resource costs	Cultural resistance to shorter hours; fear of perceived lack of dedication
Spain	Stable productivity and increased employee satisfaction in pilot programs	Sectors with high customer demands struggled with reduced hours
Belgium	Greater flexibility for employees, enhanced control over schedules	Total working hours unchanged; critics argue it doesn't address burnout or overwork effectively

Table 1: Summary of Global Experiences with a Four-Day Workweek

2.2 Taiwan's work culture and labor landscape

Taiwanese work culture is deeply rooted in long hours and a strong commitment to hard work, with three-quarters of Taiwanese employees working long hours, averaging 2,000 hours per year, significantly higher than the OECD average of 1,716 hours (Taipei Times, 2015; HR

Asia, 2023). This reflects a tradition of dedicating personal time and effort to achieve optimal results. Respect for hierarchy is another defining trait, with decisions often made in consultation with senior managers or supervisors, reflecting a strong sense of authority and order in workplaces (NNRoad, 2023). Additionally, collaboration is highly valued, as employees work effectively together to meet their goals. Taiwanese professionals also emphasize continuous learning and self-improvement, actively seeking opportunities for personal growth and showcasing a growth mindset (NNRoad, 2023). However, these positive aspects often come at the cost of personal well-being, with work-life imbalance being a pervasive issue that has significant implications for physical and mental health (HR Asia, 2023).

Taiwan has introduced several reforms aimed at improving labor conditions. According to the Labor Standards Act, weekly working hours are capped at 40 hours, and employees are entitled to two days off per week (Chung & Garcia, 2024). However, violations of labor laws remain widespread, and enforcement is inconsistent, allowing some employers to bypass regulations through loopholes or implicit expectations of unpaid overtime. In 2022, 7,348 violations of working condition laws were uncovered during inspections, with 20.44% involving unpaid overtime wages, making it the most frequently violated regulation (OSHA, 2022). Despite these findings, penalties often fail to serve as a strong deterrent, as some businesses continue to evade accountability by pressuring employees into working beyond legal limits without formal documentation. Flexible work arrangements, such as remote work, remain rare in Taiwan, with only 13% of employees having the option to work from home, compared to 37% in the United States and 28% in Japan (Chung & Garcia, 2024). This reflects Taiwan's reluctance to adapt to evolving labor trends compared to other developed nations, where flexible work models are increasingly normalized. While reforms have been introduced to improve work-life balance, significant gaps persist in policy implementation, workplace flexibility, and enforcement, indicating the need for stronger oversight and structural change.

Taiwan’s workforce continues to grapple with significant challenges, particularly regarding the effects of long working hours on occupational health. Prolonged work hours and frequent overtime have been linked to increased stress, fatigue, a higher risk of workplace injuries, and even "karoshi" (death from overwork), a phenomenon documented in several East Asian countries, including Taiwan (Jafri et al., 2024). According to research, over 30% of Taiwanese workers experience burnout, with the highest prevalence in high-pressure industries such as technology, manufacturing, and healthcare (Taipei Times, 2023). Long working hours have also contributed to a rising incidence of cardiovascular disease, with Taiwan's National Health Insurance data reporting a 20% increase in stress-related illnesses among workers in the past decade. While the government has implemented labor inspections and penalties to ensure compliance, these measures have yet to challenge the deep-rooted cultural acceptance of excessive work hours. Without a shift in workplace norms and policies, Taiwan’s workforce remains vulnerable to the long-term consequences of overwork. Younger workers (aged 20–35), in particular, are increasingly rejecting traditional expectations and advocating for improved work-life balance, as seen in recent labor protests calling for better working conditions and

stricter enforcement of labor laws (CIVICUS Monitor, 2023). Many younger employees prioritize well-being and flexibility over rigid job structures, reflecting a growing generational divide in workplace expectations. One potential solution to Taiwan's overwork culture is the implementation of a four-day workweek, which could help mitigate many of these health risks. Reducing work hours may reduce stress, while additional rest days could lower the prevalence of burnout and stress-related illnesses. Providing employees with more time for physical recovery, mental well-being, and personal activities could lead to a healthier, more productive workforce. Addressing these structural issues through labor reform and workplace innovation will be essential in ensuring sustainable work practices for Taiwan's future.

3. Methodology

This study adopts an empirical, quantitative approach to explore the feasibility and perceptions of a four-day workweek in Taiwan. The primary objective is to assess how young Taiwanese workers (aged 20-35) perceive the implications of reduced work hours and whether such a model is viable in Taiwan's labor landscape. This age group constitutes a large and dynamic segment of the national workforce. Young adults are frequently regarded as both drivers and beneficiaries of organizational innovation, especially regarding work-life balance and flexibility initiatives. According to Taiwan's Ministry of Labor, employees in this age range experience high levels of job mobility and are more likely to seek employment with progressive workplace policies. By examining their perceptions, the research aims to gauge the potential acceptance and effectiveness of a four-day workweek among those most likely to shape and be affected by future labor trends.

A structured survey was chosen as the primary research method due to its ability to collect standardized, quantifiable data efficiently. The design is particularly well-suited for capturing a snapshot of the Taiwanese people's perception and attitude toward the feasibility, benefits, and challenges of the potential four day workweek. Non-probability, convenience sampling (Saunders et al., 2007) was employed, allowing participants to voluntarily respond to the survey. Although this sampling technique may limit the generalizability of findings, it is practical for reaching a diverse set of working individuals while ensuring accessibility and ease of participation.

The primary data collection tool is a self-administered, Internet-mediated questionnaire, developed using Google Forms. The online questionnaire consists of four key sections, to systematically capture various attributes of the research questions:

1. Demographics – Collects participant background details such as age, gender, employment status, and level of education. These data enable the segmentation of responses and conducting comparisons between subgroups.
2. Awareness and Perceptions – Examines familiarity with the concept and captures their attitudes toward the idea of four-day workweek. The questions are designed to elicit both positive and negative perceptions, such as increased productivity at work vs concerns about increased workload on each workday.

3. Feasibility – Assesses perceived challenges and the practicality of implementation in Taiwan. It explores participants' views on potential barriers such as cultural resistance and corporate limits.
4. Preferences – Investigate preferred workweek structures and willingness to adopt this model. This section also includes open-ended questions that invite more detailed feedback on what changes might be necessary for facilitating this transition.

Data was collected over a four-week period. The survey link was disseminated via social media platforms (including Line and Instagram). Collected data were analyzed quantitatively, using descriptive statistics to identify trends in awareness, feasibility, and preferences. Additionally, inferential statistical methods were applied to examine associations between two or more variables. Chi-square tests of independence were selected as appropriate analytical tools due to the categorical nature of the survey data, allowing for a clear exploration of relationships between demographic factors, perceptions, and preferences. Specifically, chi-square tests of independence were used to assess relationships between categorical variables, such as education level and initial reaction to the policy. For open-ended responses, qualitative data were coded thematically, allowing for the identification of key themes and patterns in participant feedback. This approach ensures that both quantitative and qualitative insights contribute to understanding the potential adoption of a four-day workweek in Taiwan.

Ethical integrity is paramount for this study. Prior to beginning the survey, the participants were greeted with a detailed landing page message that outlined all ethical considerations of the study. It also communicated instructions regarding voluntary participation and the option to withdraw at any time without consequences. All data is confidential and used solely for academic purposes, ensuring compliance with established ethical research standards.

4. Results

This study aimed to explore perceptions regarding the potential implementation of a four-day workweek among young Taiwanese workers aged 20–35, with a particular focus on assessing the model's feasibility and attractiveness within Taiwan's labor landscape.

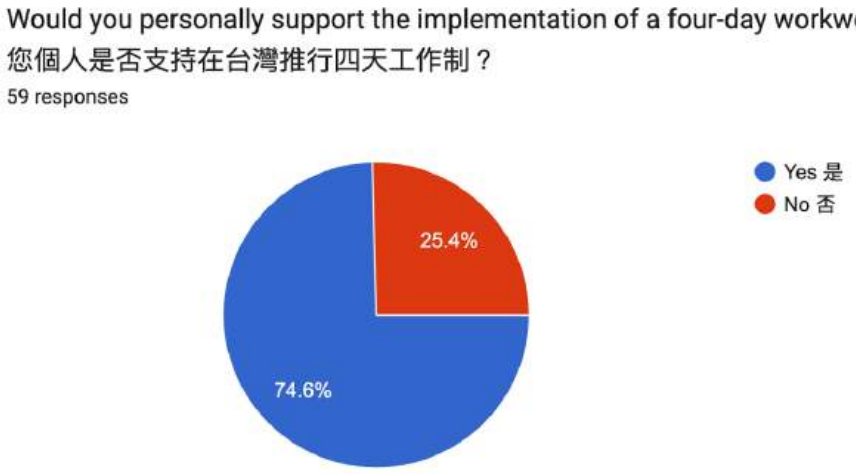


Fig 1: Perceptions about the implementation of a four-day workweek in Taiwan among respondents aged 20–35

Survey results indicate that 74.6% of respondents support the implementation of a four-day workweek in Taiwan, while 25.4% do not. This indicates a strong majority in favor, reflecting openness to alternative work arrangements among the surveyed group. However, this level of support should be interpreted with caution, as the relatively small sample size (59 participants), predominantly from the technology sectors, limits the generalizability of these findings to Taiwan's broader workforce. Still, the results suggest that a four-day workweek could gain traction if proposed as part of labor reform in Taiwan, especially among younger or more flexible segments of the workforce. It also indicates that a large number of employees may welcome changes that prioritize work-life balance. Similar trends have been observed in global case studies. For example, in the UK's four-day workweek trial, 92% of participating companies chose to continue with the policy after the trial, citing higher job satisfaction and stable productivity (Gecsoyler, 2024). Results from the UK align with the present study's results, reinforcing that reduced work hours can be popular and practical.

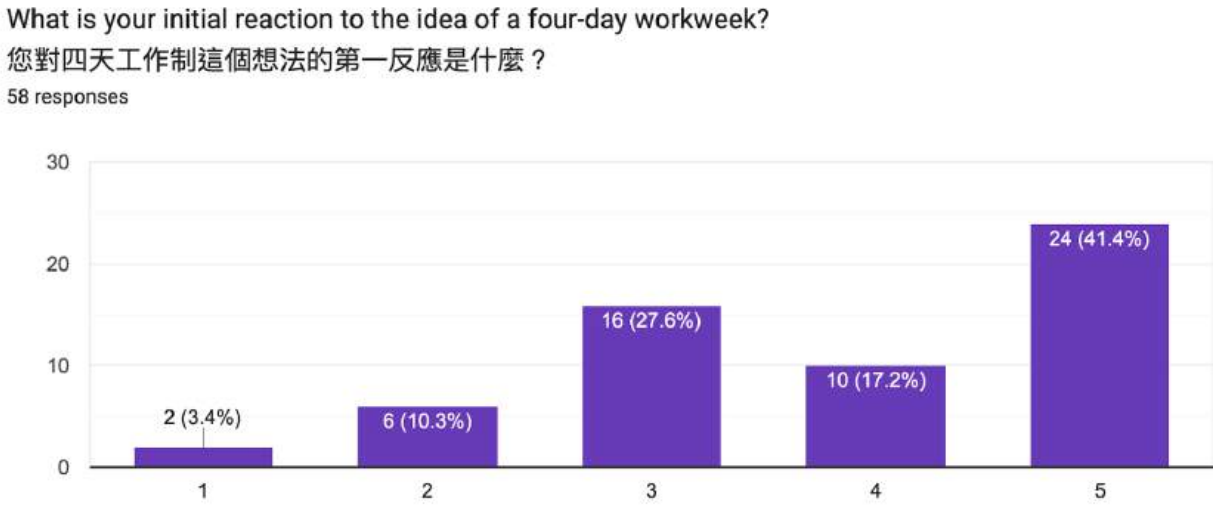


Fig 2: Distribution of participants' initial reactions to the idea of a four-day workweek

Data on participants' initial reactions, measured on a scale from 1 (negative) to 5 (positive), reveal that the majority selected either 3 (27.6%) or 5 (41.4%), while only 3.4% chose 1 and 10.3% chose 2, indicating an overall positive sentiment. This suggests a generally favorable attitude toward the idea, even before deeper practical considerations. The high share of 3 and 5 responses implies the concept resonates well with young Taiwanese workers, laying the groundwork for broader discussion and possible policy experimentation. Similar levels of initial optimism were reported in New Zealand's Perpetual Guardian trial, where employees responded enthusiastically to the change even before it began, showing anticipated work-life benefits (Hytner, 2022). This optimism was later validated by actual performance outcomes.

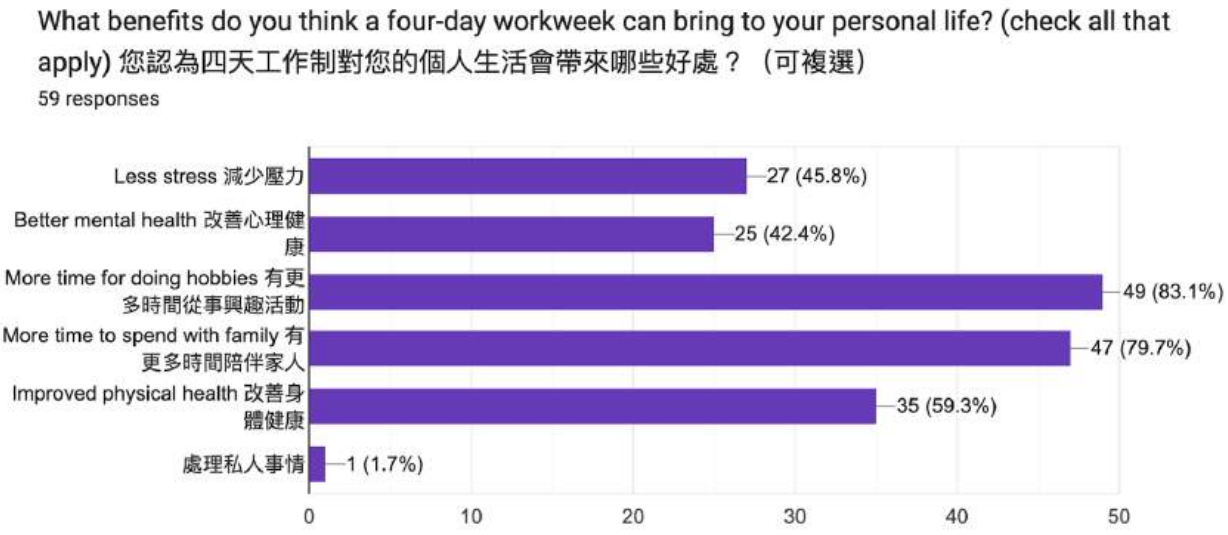


Fig 3: Perceived personal life benefits of a four-day workweek among participants

Respondents were asked to check all personal benefits they believed a four-day workweek could bring. The most selected answers were: more time for hobbies (83.1%), more time with family (79.7%), and improved physical health (59.3%). These results highlight that workers see four-day weeks as a way to reclaim time and invest in their well-being and relationships. In the UK's pilot trial, 71% of participants reported lower burnout and 54% felt better equipped to balance their lives (Lewis et al., 2023). These benefits align well with the expectations expressed in the present survey.

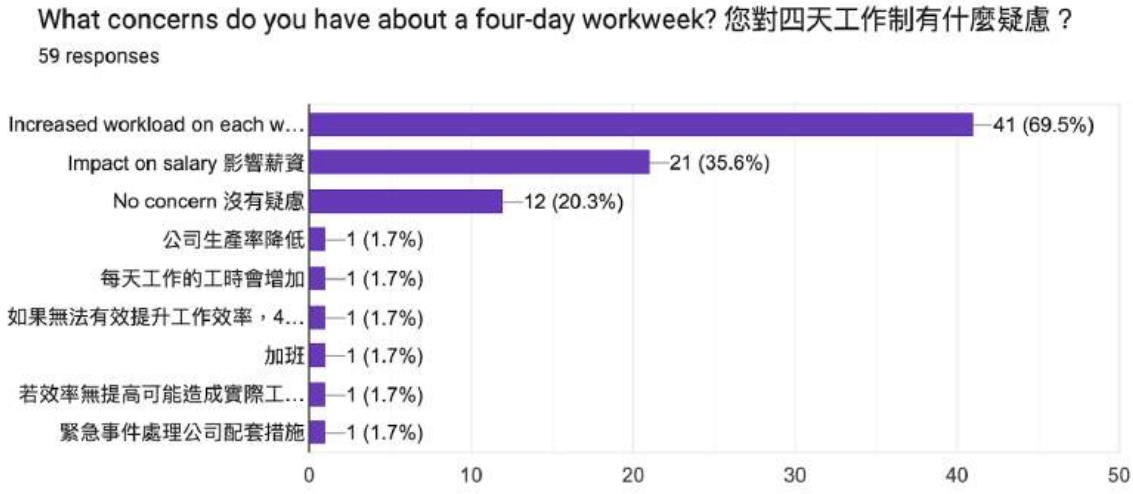


Fig 4: Participant concerns about adopting a four-day workweek in Taiwan

The top concern was increased daily workload (69.5%), followed by salary impact (35.6%). Only 20.3% reported having no concerns. While support is strong, implementation must be carefully managed to prevent the unintended consequence of simply compressing five days of work into four. Employers must address workload planning and preserve fair compensation.

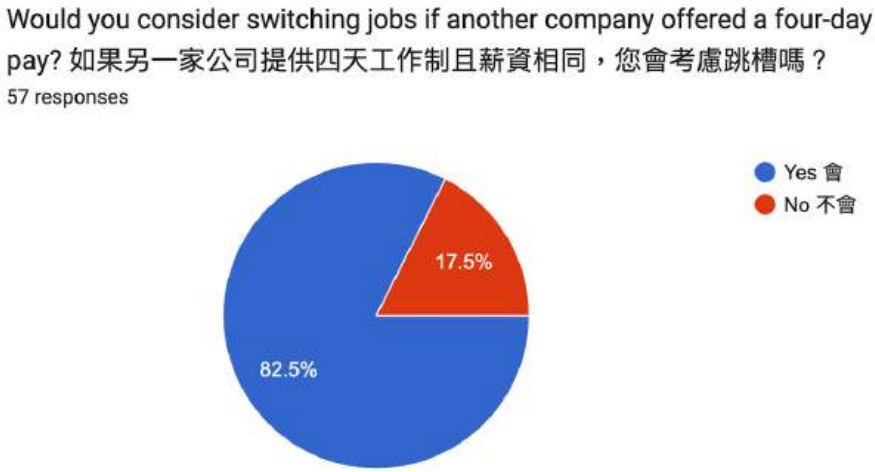


Fig 5: Willingness to switch jobs if offered a four-day workweek with equal pay

A substantial proportion, 82.5%, of respondents indicated that they would consider switching jobs if another company offered a four-day workweek with the same salary. Only 17.5% reported that they would not consider making the switch. This suggests that job seekers in Taiwan, particularly younger workers, are increasingly prioritizing work-life balance and flexibility over traditional expectations of long hours. The possibility of maintaining full pay while gaining an extra day off appears highly attractive. For employers, this represents a critical insight: offering flexible work structures could serve as a powerful tool in recruitment and retention, especially in a competitive labor market increasingly characterized by concerns over mental health and burnout. It also reflects growing dissatisfaction with the rigid five-day work structure still dominant in Taiwanese culture.

Comparative findings support this trend. In the UK four-day week pilot, participating firms reported improved retention and easier recruitment due to the appeal of the revised schedule (Lewis et al., 2023). In Japan, despite cultural resistance to fewer hours, Microsoft Japan's four-day trial led to both increased productivity and employee satisfaction, suggesting a similar appetite for reform (RoSPA, 2023). These results reflect a global convergence in workforce values, particularly among younger generations, which Taiwan is now experiencing.

5. Inferential analysis and discussion of findings

A series of chi-square tests were conducted to identify significant associations among key demographic variables, initial perceptions, feasibility beliefs, preferred work structures, and perceived government support.

5.1 Educational background and Initial reaction to the idea of a four-day workweek

A chi-square test of independence was conducted to examine the relationship between participants' highest level of education and their initial reaction to the idea of a four-day workweek. The results revealed a statistically significant association between these two variables ($\chi^2(16) = 27.1$, $p = .041$). The strength of the relationship, as measured by Cramer's V, was moderate ($V = 0.342$), aligning with Cohen's (1988) guidelines. This suggests that participants' educational background may influence their immediate perception of this alternative work model. Notably, the post-hoc tests revealed that individuals holding doctoral degrees were the most likely to rate their initial reaction as 1 (negative), indicating a higher level of skepticism or critical evaluation toward the four-day workweek. One possible explanation is that doctoral degree holders may occupy positions with high levels of responsibility, autonomy, or research obligations, making them particularly cautious about altering established work patterns. These individuals might also perceive shorter workweeks as potentially disruptive to productivity or incompatible with their professional expectations, especially in academic or highly specialized fields. Conversely, participants with lower educational qualifications might view reduced hours more favorably, possibly driven by aspirations for better work-life balance or fewer perceived disruptions to their routine tasks. These findings suggest a meaningful link between educational attainment and openness to innovative work arrangements, possibly due to differences in job responsibilities or industry norms.

5.2 Initial reactions and Perceived feasibility of a four-day workweek

A statistically significant relationship was found between participants' initial reaction to the idea of a four-day workweek and their belief about whether it is feasible in Taiwan's economic and cultural context ($\chi^2(4) = 20.2$, $p < .001$). The strength of this relationship, as measured by Cramer's $V = 0.601$, indicates a strong association (Cohen, 1988). Those who gave the highest reaction rating of 5 (very positive) were most likely to believe the policy is feasible, while those who rated their initial reaction as 3 or lower tended to believe it is not. Post hoc residuals showed that the strongest positive contributors to the significance were the "5–Yes" (residual = 2.403) and "3–Yes" (residual = 2.179) cells. In contrast, positive reactions were significantly underrepresented among those who thought the policy was unfeasible. This suggests that first impressions play a crucial role in shaping perceptions of feasibility, with optimistic or favorable respondents much more likely to believe in the workability of a four-day system within Taiwan. Those who initially perceive the four-day workweek positively might interpret potential barriers—such as employer resistance, workload redistribution, or sector-specific challenges—as manageable. In contrast, individuals with initial reservations

might focus on perceived risks or practical hurdles, such as economic disruption or cultural resistance.

5.3 Feasibility and Preferred structure for a four-day workweek

A statistically significant relationship was found between participants' perceived feasibility of implementing a four-day workweek in their industry and their preferred structure for such a workweek ($\chi^2(12) = 28.6$, $p = .004$). This result suggests that perceptions of practicality are linked to how individuals imagine or desire the new work model to be implemented. The strength of this association, as indicated by Cramer's $V = 0.428$, reflects a moderate effect size (Cohen, 1988), pointing to a notable connection between feasibility and structural preference. Post hoc residual analysis showed that those who rated feasibility as very low (1) were significantly more likely to prefer the traditional five-day workweek (residual = 2.287), and less likely to prefer flexible scheduling (residual = -2.2787). On the other hand, those in the same low-feasibility group showed an unexpectedly high preference for the "32-hour four-day model" (residual = 2.219). This mix of conservative and progressive preferences suggests that while some respondents may view reduced-hour workweeks as ideal in theory, they may still doubt their practicality in their current industry setting. This reflects a noteworthy tension within the Taiwanese workforce—while there is conceptual openness to alternative work schedules that promise improved work-life balance and reduced stress, ongoing uncertainty about how such models can be realistically implemented across different sectors may limit full endorsement.

5.4 Government support and Perceived workplace benefits

A statistically significant relationship was found between participants' perceptions of government support and the perceived benefits of a four-day workweek in their work life ($\chi^2(44) = 81.2$, $p < .001$). The strength of this association, indicated by Cramer's $V = 0.602$, suggests a strong relationship between these two categorical variables (Cohen, 1988). Post hoc Pearson residuals revealed one key contributor to this association: participants who rated government willingness and capability as 5 (very high) were significantly more likely to select "Enhanced creativity" as a benefit (residual = 7.35). This suggests that those who are most optimistic about governmental involvement are also those who see creativity—often associated with flexible working conditions, autonomy and a healthier work-life balance—as one of the most valuable outcomes of a shorter workweek in the workplace. Other benefits such as reduced burnout or job satisfaction were not found to have significant associations in this residual analysis. The implications of this result are noteworthy for policymakers. It appears that perceived government support may shape how individuals view the potential of the four-day workweek—not just in terms of feasibility, but also in terms of outcomes. Those optimistic workers are more likely to envision transformative, innovation-driven benefits from workplace reforms.

5.5. Experience of flexible work arrangements and Belief in the feasibility of a four-day workweek

A statistically significant relationship was found between participants' experience with flexible or reduced working hours and their belief in the feasibility of implementing a four-day workweek in Taiwan ($\chi^2(1) = 4.45$, $p = 0.035$). The strength of this association, measured by Cramer's $V = 0.279$, indicates a weak to moderate effect size (Cohen, 1988). Participants who had prior experience with flexible or reduced hours were more likely to believe that a four-day workweek is feasible, whereas those without such experience were more likely to express doubts about its practicality. Although post hoc residuals did not exceed the threshold for statistical significance ($|residual| > 2$), the distribution pattern reinforces the trend: among those who had never experienced flexible work, there was an underrepresentation of positive feasibility ratings (residual = -1.27), and among those who had, there was an overrepresentation (residual = 1.13). These results suggest that exposure to alternative work arrangements may shape more optimistic attitudes toward work reform.

6. Insights from Open-Ended Responses

The qualitative responses complement and expand upon the quantitative findings, particularly regarding skepticism about the feasibility of a four-day workweek. They highlight several critical conditions for successful implementation: strong and transparent leadership, clearly defined roles and responsibilities, robust political and institutional support, and careful planning to ensure fair and manageable workloads. Without these structural supports, participants expressed concerns that the four-day model may exacerbate rather than alleviate pressure, especially in traditional industries workplaces. These insights reinforce the need for systemic changes beyond policy proposals, including cultural transformation at both organizational and national levels.

- Many respondents emphasized that a supportive employer mindset is crucial. Without buy-in from company leaders, especially in SMEs, structural changes like a four-day workweek are unlikely to succeed.
- Respondents noted the need for mutual trust; employers must trust employees to work efficiently, and employees must commit to completing tasks within reduced hours.
- There is a strong call for government-led policy reforms, including legal protections, enforcement mechanisms, and public awareness campaigns to ensure implementation and compliance.
- Suggestions include better task delegation, clearer role definition, and improved time management tools (e.g., reducing unnecessary meetings) to help teams adapt to shorter workweeks.
- Some responses stressed that digital tools, automation, and AI integration are key to maintaining productivity in a four-day model, especially in manufacturing or traditional industries.

- A few respondents expressed worries that if supervisors' mindsets don't change, reduced hours may result in increased blame or pressure on workers rather than systemic improvement.

7. Comparative analysis of emerging trends and policy implications

The findings of the present study align with existing global literature while contributing novel insights specific to Taiwan's cultural and labor context. This section compares the five emerging trends identified through the data with established research and outlines their implications for policymaking. Overall, the evidence supports the need for carefully framed, context-sensitive approaches to labor reform in Taiwan, particularly regarding the implementation of a four-day workweek.

Trend 1: Higher Education and Initial Skepticism

One of the most prominent findings was that doctoral degree holders were the most likely to react negatively to the idea of a four-day workweek. This observation is consistent with Golden and Wiens-Tuers (2006), who argued that individuals with higher education levels often work longer hours and shoulder more responsibilities, which may lead to greater skepticism toward changes that threaten existing work structures. These individuals may also operate in specialized or research-intensive roles where conventional measures of productivity are harder to reconfigure. While this finding confirms existing research, the particularly strong doubt observed among doctoral degree holders adds depth to the current literature, representing a contribution to our understanding of how professional status and role complexity influence openness to labor reforms. For policymakers, this suggests that communication strategies must address not just general awareness, but also the specific concerns of professionals in high-responsibility or autonomous roles.

Trend 2: Initial Reactions Shape Feasibility Beliefs

The data revealed a strong association between participants' initial reactions and their perceived feasibility of the four-day workweek. Those who reacted positively were more likely to consider it workable in Taiwan's economic and cultural context. This aligns with existing literature on organizational change, which emphasizes the power of initial attitudes in shaping future acceptance (Kotter & Schlesinger, 2008). Moreover, as Haar et al. (2019) note, individuals who anticipate positive work-life balance outcomes are more receptive to change. Similarly, Baker, Avery, and Crawford (2007) found that professionals who had optimistic expectations about remote work reported higher satisfaction and greater long-term acceptance of the arrangement. This supports the idea that initial perceptions play a crucial role in shaping engagement with workplace reforms more broadly. For policymakers, this highlights the importance of early framing and public education. Building awareness of successful case studies, offering evidence-based reassurances, and clearly communicating implementation strategies can positively influence initial perceptions, ultimately shaping the public's perceived feasibility. For

Taiwan, a society historically characterised by long hours and intense workloads, cultivating positive initial impressions of shorter workweeks could play a pivotal role in enhancing public belief in their practicality.

Trend 3: Mixed Preferences Among Those Doubting Feasibility

A particularly nuanced insight emerged from respondents who doubted the feasibility of a four-day workweek in their own industry but still expressed a preference for progressive models like the 32-hour week. This tension between aspiration and skepticism mirrors what Haraldsson and Kellam (2021) describe as an "implementation paradox," where workers are drawn to the idea of reduced hours yet uncertain about its viability in practice. Similarly, a 2021 study by the Social Market Foundation in the UK found that although enthusiasm for a shorter workweek was especially high among higher earners and professionals, only 11% were willing to take a pay cut to achieve it, highlighting the gap between abstract support and practical acceptance (Baker et al., 2021). Further reinforcing this pattern, researchers from What Works Wellbeing (2023) observed persistent skepticism among UK employees, even as four-day week trials produced overwhelmingly positive outcomes. This indicates that doubts can persist not because of poor results, but due to ingrained caution about long-term implementation or trade-offs. For Taiwan, where long-standing labor norms and high-pressure work environments persist, this finding underscores the importance of gradual implementation. Launching industry-specific pilot programs and providing tangible demonstrations of success could help bridge the gap between concept and confidence. By offering multiple structural options, such as flexible schedules or reduced-hour models without compressed days, policymakers can accommodate the diverse needs and readiness levels within the workforce. This measured approach may be key to transforming cautious curiosity into actionable support.

Trend 4: Perceived Government Support Shapes Expected Benefits

Participants who believed the government was both willing and capable of supporting labor reform were significantly more likely to associate a four-day workweek with transformative benefits such as enhanced creativity. This suggests that institutional trust plays a key role in shaping expectations of workplace innovation. Similar findings appear in international public policy literature, where perceived government involvement increases public optimism about reform outcomes (Moynihan & Landuyt, 2009). Conversely, when citizens doubt governmental competence or willingness, they are significantly less likely to expect meaningful reforms to succeed. Data from the OECD's 2024 survey on trust in public institutions show that policy skepticism correlates with lower expected benefits, even when policies are well-designed and evidence-based (OECD, 2024). This indicates that trust not only shapes attitudes toward feasibility, but also influences whether people believe reforms will produce meaningful or creative workplace improvements. For Taiwan, this means that government messaging and visible commitment, through legislation, incentives, or strategic partnerships, could enhance perceived legitimacy and shift public sentiment from cautious to hopeful.

Trend 5: Prior Flexible Work Experience Linked to Optimism

Respondents with previous flexible work or reduced-hour experience were significantly more likely to believe that a four-day workweek is feasible. This finding aligns with research by Shockley et al. (2017), who argue that personal exposure to flexible work structures builds psychological readiness for change and increases openness to future innovations. For Taiwan, this implies that even limited exposure to flexible arrangements, such as hybrid models or flexible start/end times, could gradually foster a culture of acceptance. Employers and policymakers can leverage this by initiating small-scale pilots that familiarize workers with alternative scheduling and build the confidence needed to support broader reforms.

8. Broader implications for Taiwan's labor landscape

The survey reveals growing support among young Taiwanese workers for adopting a four-day workweek, with strong interest driven by benefits such as improved well-being, reduced burnout, and better work-life balance. A majority of respondents said they would support the implementation or even switch jobs for such a model. However, skepticism remains, particularly among those unfamiliar with flexible work arrangements or working in industries perceived as less adaptable. Importantly, belief in government support was strongly linked to optimism about the policy's outcomes, especially enhanced creativity. This highlights the critical role of public trust and leadership in driving workplace reform. For policymakers, the path forward may lie in gradual implementation via pilot programs, flexible models, and industry-specific adaptations that build confidence and demonstrate real-world feasibility. These steps could pave the way for Taiwan to modernize its labor landscape while supporting both productivity and employee well-being.

9. Conclusion

This study set out to explore how young Taiwanese workers (aged 20 - 35) perceive the feasibility and desirability of a four-day workweek, and to identify the factors that most strongly shape those perceptions. Drawing on a nationwide online survey and a series of chi-square tests, four clear findings emerge. These findings confirm and extend existing international literature, notably research by Kotter and Schlesinger (2008) and Haar et al. (2019), which highlight the significance of initial perceptions and institutional support in successfully communicating plans for, and managing, organisational change.

First, education matters. The findings indicate that educational background plays a significant role in shaping initial responses to the four-day workweek, with those holding doctoral degrees exhibiting greater skepticism, possibly due to higher job responsibility or incompatibility with established professional norms. Second, initial impressions count. Participants with favorable first impressions were substantially more likely to view the four-day model as feasible, reinforcing the importance of early framing and public communication in shaping perceptions of reform. Third, perceived practicality and preferred workweek structure

are interrelated. Even among those who doubted the feasibility, there was notable support for progressive formats such as the 32-hour model, suggesting aspirational openness tempered by realistic concerns. Finally, trust in government proved to be a decisive factor. Respondents who believed in the government's willingness and capacity to implement reform were more likely to associate the four-day workweek with innovation-related outcomes.

Taken together, these insights broadly suggest that effective implementation of a four-day workweek in Taiwan will require strong government leadership, targeted public communication, and adaptive policy design. From a practical standpoint, policy-makers should introduce gradual pilot programs, while pairing flexible scheduling options with evidence-based case studies to help turn public interest into real action by easing concerns from employers and institutions.

While these findings offer valuable contributions, the study has its own limitations. The use of convenience sampling limits generalizability, and given that a majority of respondents work in technology-driven industries, the results may reflect sector-specific flexibility and openness to alternative work models, perspectives that may differ significantly from those in more traditional sectors such as manufacturing. Future research should adopt probability sampling and include industry-specific studies to better reflect the diversity of Taiwan's labor market. It should also consider longitudinal designs to examine how employee perceptions evolve over time, particularly before and after implementation, and conduct sector-specific analyses to identify both the challenges and opportunities faced by different industries in adapting to a four-day workweek model.

Works Cited

Baker, E., Avery, G. C., & Crawford, J. (2007). Satisfaction and perceived productivity when professionals work from home. Research and Practice in Human Resource Management, 15(1), 37–62.

Baker, R., Gardiner, L., & Valero, A. (2021). A question of time: Work and productivity growth in the UK. Social Market Foundation. https://www.smf.co.uk/wp-content/uploads/2021/07/A-question-of-time-July-2021.pdf

Chung L., Garcia S. (2024). The Labor Standards Act limits working hours in Taiwan. Taipei Times. https://www.taipeitimes.com/News/taiwan/archives/2024/12/13/2003828450

CIVICUS Monitor. (2023). Taiwan: Multiple protests call for improved labor conditions ahead of elections. CIVICUS. https://monitor.civicus.org/explore/taiwan-multiple-protests-call-for-improved-labour-conditions-ahead-of-elections/

Gecsoyler, S. (2024, August 30). This article is more than 3 months old: Workers in the UK could get right to request four-day compressed week. The Guardian. https://www.theguardian.com/business/article/2024/aug/30/workers-uk-four-day-week-compressed-hours-flexible-working

Golden, L., & Wiens‐Tuers, B. (2006). To your happiness? Extra hours of labor supply and worker well‐being. Journal of Socio‐Economics, 35(2), 382–397. https://doi.org/10.1016/j.socec.2005.11.039

Haar, J. M., Russo, M., Suñe, A., & Ollier-Malaterre, A. (2019). Outcomes of work-life balance on job satisfaction, life satisfaction, and mental health: A study across seven cultures. Journal of Vocational Behavior, 112, 235-247.
https://www.sciencedirect.com/science/article/abs/pii/S0001879114001110

Haraldsson, G. D., & Kellam, J. (2021). Going public: Iceland's journey to a shorter working week. Reykjavík & London. https://autonomy.work/wp-content/uploads/2021/06/ICELAND_4DW.pdf

HR Asia. (2023). Taiwan has the sixth-longest work hours. HR Asia. https://hr.asia/featured-news/taiwan-has-sixth-longest-work-hours/

Hurst, L. (2023). Valencia's four-day week trial results in less stress and more sleep. Euronews. https://www.euronews.com/next/2023/10/31/workers-in-valencias-4-day-week-trial-report-less-stress-and-more-socialising

Hytner, M. (2022). Australian workers at global company to trial four-day week after success of New Zealand pilot. The Guardian. https://www.theguardian.com/australia-news/2022/nov/02/australian-workers-at-global-company-to-trial-four-day-week-after-success-of-new-zealand-pilot

Kassam, A. (2021). Spain to launch a trial of four-day working week. The Guardian. https://www.theguardian.com/world/2021/mar/15/spain-to-launch-trial-of-four-day-working-week

Kobie, N. (2024). Iceland's four-day working week trials have been a roaring success – economic growth spiked, workers were happier, and burnout plummeted. ITPro. https://www.itpro.com/business/business-strategy/icelands-four-day-working-week-trials-

have-been-a-roaring-success-economic-growth-spiked-workers-were-happier-and-burnout-plummeted
Kotter, J. P., & Schlesinger, L. A. (2008). Choosing strategies for change. Harvard Business Review, 86(7/8), 130-139.
Lee I. (2023). The growing case for change: Labor protests and workplace reforms in Taiwan. Taipei Times. https://www.taipeitimes.com/News/taiwan/archives/2023/10/23/2003808095
Ministry of Justice, Taiwan. (n.d.). Labor Standards Act – Work Hours Regulations. Laws & Regulations Database of the Republic of China. https://law.moj.gov.tw/ENG/LawClass/LawParaDeatil.aspx?pcode=N0030001&bp=4
Ministry of Labor, Taiwan. (2023). Minimum wage increases and labor rights protections. https://www.mol.gov.tw
Moynihan, D. P., & Landuyt, N. (2009). How do public organizations learn? Bridging cultural and structural perspectives. Public Administration Review, 69(6), 1097–1105.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3492431
NNRoad. (2023). Taiwan's work culture. NNRoad. https://nnroad.com/blog/taiwan-work-culture/
OSHA, Taiwan. (2022). Annual report of labor inspection statistics. https://www.osha.gov.tw/media/qd0a5otd/annual-report-of-labor-inspection-statistics-in-2022.pdf
RoSPA. (2023). Four-day workweek: Trial results and implications. Royal Society for the Prevention of Accidents. https://www.rospa.com/getdoc/3d130499-25ad-4b1d-8040-2615e7cd6c6f/four-day-week
Saunders M., Lewis P., Thornhill A. (2007). Research Methods for Business Students. https://www.researchgate.net/publication/240218229_Research_Methods_for_Business_Students
Shockley, K. M., Shen, W., DeNunzio, M. M., Arvan, M. L., & Knudsen, E. A. (2017). Disentangling the relationship between flexible work arrangements and work–family conflict: A meta-analysis. Journal of Applied Psychology, 102(2), 241–255.
https://psycnet.apa.org/record/2017-32286-001
Van Olmen, C. (2022). Belgium Labour Deal 2022: 4-day work week and alternating work week. L&E Global. https://leglobal.law/2022/11/27/belgium-labour-deal-2022-4-day-work-week-and-alternating-work-week/
Vega, G. (2022). How two companies in Spain moved to a four-day week without cutting salaries. El País. https://english.elpais.com/science-tech/2022-01-20/how-two-companies-in-spain-moved-to-a-four-day-week-without-cutting-salaries.html
What Works Wellbeing. (2023). Findings from the four-day work week trial. https://whatworkswellbeing.org/blog/findings-from-four-day-work-week-trial/
Wong, K., Chan, A. H., & Ngan, S. C. (2019). The effect of long working hours and overtime on occupational health: A meta-analysis of evidence from 1998 to 2018. International Journal of Environmental Research and Public Health, 16(12), 2102. https://pmc.ncbi.nlm.nih.gov/articles/PMC6617405/

Adolescent Cannabis Use: Effects on Brain Development and Applications to Schools and Communities By Prajna Sakshi

Abstract

Cannabis, a mind-altering drug, has grown to be one of the most commonly abused drugs by teens. The following review article summarizes how cannabis use alters the developing adolescent brain and how it can relate to emotional and mental dysregulation. The article discusses how we can use this knowledge to strengthen communities and improve school systems. Schools and communities can apply this knowledge effectively by educating students on the biological aspect and understanding cannabis use among teens around the world.

Introduction

Cannabis, a mind-altering drug also known as marijuana, pot, weed, and chronic, has grown to be one of the most commonly abused drugs by teens [1]. A study found that in 2022, 30.7% of US 12th graders reported ever having used cannabis [2]. Adolescent use of cannabis has remained steady over the last three decades, and vaping has emerged as an appealing way for teens to use the drug [3][4]. Existing research has demonstrated that adolescent brains are particularly susceptible to the effects of cannabis use, negatively affecting mental and physical health in teens [2].

Cannabis and the Adolescent Brain: Adolescent Brain Development Processes

Adolescence is a pivotal time for brain development, particularly in the prefrontal cortex (PFC) which is responsible for regulating thoughts, emotions, and actions [5]. The striatum (STR), tightly connected to the PFC, is involved in motor control, reward processes, and executive functions such as decision-making [6][7]. The connection between the PFC and striatum is said to be critical for cognitive processes and is thought to be strengthened with age, as shown in Figure 1 [8]. These brain structures experience a multitude of developmental changes throughout adolescence. Interestingly, adolescence is a period associated with an increase in substance use. The prevalence of substance use during adolescence is likely influenced by factors such as stress, emotional regulation, and reward processing, as the brain structures responsible for these functions are not yet fully developed [9].

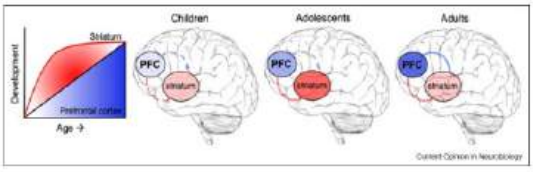


Figure 1. Brain model of the prefrontal cortex (PFC) and ventral striatal cortex throughout stages of development. Solid lines demonstrate mature connection while dotted lines indicate immaturity. Pigmented colors indicate better signaling [8].

How Cannabis Influences the Brain

Cannabinoids are compounds found naturally in cannabis, with the two primary types being tetrahydrocannabinol (THC) and cannabidiol (CBD). THC is the more psychoactive

component. Cannabis acts on the adolescent brain through cannabinoid receptors, which are present throughout the body. During times of stress, the brain releases its own cannabinoids called endocannabinoids, located in the brain, muscles, and other organs and tissues [10][11]. CB1, the first type of cannabinoid receptor, controls striatal dopamine, a chemical messenger in the brain associated with feelings of motivation and pleasure [12]. The second cannabinoid receptor (CB2) is located in the brain and periphery immune systems, and helps control antinociception, or pain relief, supporting marijuana's medicinal properties. THC binds to endocannabinoid receptors in the brain, activating the endocannabinoid system. THC's high binding affinity alters how the receptors function [13]. For instance, in CB1, the binding of THC increases dopamine release. Repetitive use causes a reduced dopamine response, known as the blunting of the dopamine system, aiding in addiction [14]. Ultimately, these modifications of the adolescent endocannabinoid system can result in inferior cognitive processes in adulthood [15].

Moreover, cannabis impacts the brain during adolescence through the PFC. Through research and experiments conducted on animals such as rats, scientists have concluded that THC exposure postpones the development of the PFC in the adolescent brain. For example, rats exposed to THC performed inferiorly than the control animals in a maze learning task [5]. Recurrent cannabis use is correlated to reduced gyrification in the developing PFC, which indicates a decrease in folding pattern of the cerebral cortex. This folding process, referred to as cortical gyrification, increases the cortex's surface area and number of neurons it contains, ultimately affecting development [16].

Additionally, gamma-aminobutyric acid (GABA) is the brain's main inhibitory neurotransmitter, or chemical messenger allowing neuron communication, in the central nervous system (CNS). Inhibitory neurotransmitters decrease the nerve cells' ability to send, receive, and create chemical messages to other nerve cells, therefore blocking certain signals in the CNS. GABA is thought to contribute to producing calming effects and controlling stress and anxiety [17]. Furthermore, glutamate, an abundant excitatory neurotransmitter in the CNS, is critical for proper brain function, and plays a role in memory and learning. Excitatory neurotransmitters stimulate nerve cells, increasing the likelihood of the chemical message moving along nerve cells [18]. Cannabis use can result in GABA synthesis reduction, which is linked with conditions such as schizophrenia and major depressive disorder. In adolescents, constant THC exposure results in GABA insufficient function (hypofunction), in the adult PFC [19]. Additionally, a 2019 study reported that a single dose of THC resulted in increased striatal glutamate levels. This increase is associated with high risks of psychosis, symptoms that occur when an individual has trouble recognizing what is real and what is not [20].

Emotional and Mental Dysregulation

THC causes an imbalance in the production of neurotransmitters glutamate and GABA, which in turn disrupts the ratio between excitatory (E) and inhibitory (I) neurons in the brain, known as the E:I ratio [21]. This imbalance can lead to cognitive impairment and brain dysfunction, linking to emotional and mental dysregulation [22]. Emotional regulation, crucial

for the counterbalance of negative feelings, is an individual's ability to control their emotions. The incorporation of cannabis into this fragile system, especially in the adolescent brain, may lead to it becoming dysfunctional. Studies have shown that emotional dysregulation is a transdiagnostic risk factor for substance use. Transdiagnostic risk factors are disrupted processes such as emotions or cognitions that increase the likelihood of developing psychiatric disorders [23]. For instance, a study from Columbia University suggested that teen cannabis users are two to four times more likely to develop a psychiatric disorder in their life. Cannabis users may experience short term advantages including relaxation, pleasure, and amelioration of stressful emotions. However, these temporary effects only fortify the addiction, resulting in an increase in emotional dysregulation and a heavier detriment on mental and physical health [24].

Application to Communities: Government Policies on Cannabis

There has been debate regarding the positive versus negative effects recreational marijuana legalization has on communities; some argue that cannabis use is widespread among young adults and causes less harm compared to alcohol and opioids, making legalization less of a concern. Furthermore, they argue that legalizing marijuana is preferable because it abolishes illegal sales and reduces adolescent access. However, the acceptance of marijuana may grow to be a larger problem, as it could lead to an increase in users and higher rates of addiction. Additionally, lower cannabis prices may result in easy access, potentially driving widespread use [25].

Cannabis Use Globally

In high-income countries such as North America, Europe, and Oceania, there is a greater prevalence of cannabis use compared to that in middle- and lower-income countries [26]. The variance of cannabis use across countries are affected by factors including religion, politics, and policies. For example, adolescents in Asian, African, and European countries have been reported to use cannabis more than the senior population, perhaps due to the growing acceptance of cannabis in society. A common finding was that cannabis use is more common in individuals receiving inadequate support from others or for those who have family history. The higher acceptance a country has towards cannabis, the more likely individuals are to use it.

Decreasing Cannabis Use in Adolescents

Although countries have varying standpoints on cannabis legalization, many are attempting to raise awareness to adolescents and the public and incorporate prevention programs. Studies have shown that school-based prevention programs are one of the most effective methods for decreasing substance use among youth [26]. However, ineffective implementation can significantly undermine the success of these programs. Programs in which skill-based training is executed tend to yield better results. For instance, the EU-Dap Trial, which evaluated European Unplugged school programs that utilized interactive activities and skill-building exercises, found the program effective in reducing cannabis use, cigarette smoking, and episodes of drunkenness.

These positive effects were sustained at a 15-month follow-up, particularly among adolescents in schools with average to low socioeconomic levels [27]. Additionally, a study involving 142 adolescents in Spain evaluated the impact of the Reasoning and Rehabilitation V2 Program (R&R2), a cognitive-behavioral intervention designed to prevent substance use among students at risk of academic failure. Participants were randomly assigned to either the experimental group, which received the intervention, or a control group. The R&R2 program, divided into modules that focused on a specific skill set, employed a combination of techniques, such as stress management, impulse control, group discussions, habit replacement, and goal setting, among others, to promote lasting behavioral change. A post-test survey revealed that those in the experimental group experienced a notable reduction in cannabis use, as well as a decline in cigarette smoking and alcohol consumption [28]. In regard to spreading awareness, school programs in Nepal, Spain, Malaysia, and France use unique ways such as film contests and digital tools [29]. Additionally, schools could benefit from integrating peer mentorship programs and hosting open forums where teens can discuss their challenges and pressures in a safe, supportive setting. Identifying major stressors in teens, educating students on healthy coping skills, and informing students of real-life examples of the effects of cannabis can play a role in reducing cannabis dependence.

Conclusion

The high prevalence of cannabis use among adolescents, along with its detrimental effects such as addiction, long-term brain damage, and emotional and mental dysregulation, highlights the crucial need for targeted interventions. The existing research presented regarding the adolescent brain and effective strategies for reducing cannabis abuse in teens, particularly within educational settings, enables us to identify key areas for action and community engagement. Being able to address and solve these issues proactively will assist communities in reducing drug use in the young population and contributing to healthier developmental outcomes.

Works Cited

[1] (04/01/2023). 2021 NSDUH Annual National Report. Substance Abuse and Mental Health Services Administration. https://www.samhsa.gov/data/report/2021-nsduh-annual-national-report

[2] (15/02/2023). Cannabis and Teens. Centers for Disease Control and Prevention. https://www.cdc.gov/cannabis/health-effects/cannabis-and-teens.html

[3] (24/12/2019). Cannabis (Marijuana) DrugFacts. National Institute on Drug Abuse. https://nida.nih.gov/publications/drugfacts/cannabis-marijuana

[4] Keyes, Katherine M. et al. (20/09/2022). Temporal trends in alcohol, cannabis and simultaneous use among 12th-grade U.S. adolescents from 2000 to 2020: Differences by sex, parental education, and race and ethnicity. Alcoholism: Clinical and Experimental Research. https://doi.org/10.1111/acer.14914

[5] Shen, Helen. (2020). Cannabis and the adolescent brain. Proceedings of the National Academy of Sciences. https://www.pnas.org/doi/epdf/10.1073/pnas.1920325116

[6] Zhang, Yan et al. (2016). Functional connectivity between prefrontal cortex and striatum estimated by phase locking value. Cognitive Neurodynamics. 10.1007/s11571-016-9376-2

[7] Helion, Chelsea et al. (2019). Ventral Striatum. Handbook of Clinical Neurology. https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/ventral-striatum

[8] Current Opinion in Neurobiology. (2010). https://www.ncbi.nlm.nih.gov/core/lw/2.0/html/tileshop_pmc/tileshop_pmc_inline.html?title=Click%20on%20image%20to%20zoom&p=PMC3&id=3014528_nihms174656f1.jpg

[9] Hammond, Christopher J. et al. (2014). Neurobiology of Adolescent Substance Use and Addictive Behaviors: Prevention and Treatment Implications. Adolescent Medicine: State of the Art Reviews. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4446977/

[10] Paul, Marla. (21/09/2023). Your body's own cannabinoid molecules calm you during stress. Northwestern Now. https://news.northwestern.edu/stories/2023/09/your-bodys-own-cannabinoid-molecules-calm-you-during-stress/#:~:text=When%20you%20are%20under%20stress,molecules%20were%20not%20well%20known.

[11] Pietro, De MaryAnn. (27/02/2021). What to know about endocannabinoids and the endocannabinoid system. MedicalNewsToday. https://www.medicalnewstoday.com/articles/endocannabinoid

[12] Watson, Stephanie. (18/04/2024). Dopamine: The pathway to pleasure. Harvard Health Publishing. https://www.health.harvard.edu/mind-and-mood/dopamine-the-pathway-to-pleasure

[13] Vuckovic, Sonja et al. (12/11/2018). Cannabinoids and Pain: New Insights From Old Molecules. Frontiers in Pharmacology Section Neuropharmacology.

https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2018.01259/full
[14] Bloomfield A P Michael et al. (16/5/2017). The effects of Δ9-tetrahydrocannabinol on the dopamine system. Nature. 10.1038/nature20153
[15] Jacobus, Joanna & Susan F. Tapert. (2014). Effects of Cannabis on the Adolescent Brain. Current Pharmaceutical Design. https://doi.org/10.2174/13816128113199990426
[16] Shollenbarger, Skyler G. et al. (23/07/2015). Impact of cannabis use on prefrontal and parietal cortex gyrification and surface area in adolescents and emerging adults. Developmental Cognitive Neuroscience. 10.1016/j.dcn.2015.07.004
[17] Gamma-Aminobutyric Acid (GABA). Cleveland Clinic. https://my.clevelandclinic.org/health/articles/22857-gamma-aminobutyric-acid-gaba
[18] Glutamate. Cleveland Clinic. https://my.clevelandclinic.org/health/articles/22839-glutamate
[19] Renard, Justine et al. (01/07/2018). Effects of Adolescent THC Exposure on the Prefrontal GABAergic System: Implications for Schizophrenia-Related Psychopathology. Frontiers in Child and Adolescent Psychiatry. https://doi.org/10.3389/fpsyt.2018.00281
[20] Colizzi, Marco et al. (15/02/2019). Delta-9-tetrahydrocannabinol increases striatal glutamate levels in healthy individuals: implications for psychosis. Molecular Psychiatry. https://doi.org/10.1038/s41380-019-0374-8
[21] Alreja, Arish, Ilya Nemenman, Christopher J. Rozell. (21/01/2022). Constrained brain volume in an efficient coding model explains the fraction of excitatory and inhibitory neurons in sensory cortices. PLOS Computational Biology. https://doi.org/10.1371/journal.pcbi.1009642
[22] Maestu, Fernando et al. (2021). Neuronal excitation/inhibition imbalance: core element of a translational perspective on Alzheimer pathophysiology. Ageing Research Reviews. https://doi.org/10.1016/j.arr.2021.101372
[23] Cavalli, Jessica M, Cservenka Anita. (2020). Emotion Dysregulation Moderates the Association Between Stress and Problematic Cannabis Use. Frontiers Psychiatry. 10.3389/fpsyt.2020.597789
[24] De Berardis, Domenico et al. (26/08/2020). Emotional Dysregulation in Adolescents: Implications for the Development of Severe Psychiatric Disorders, Substance Abuse, and Suicidal Ideation and Behaviors. Brain Sciences. https://www.mdpi.com/2076-3425/9/4/76.
[25] Hall, Wayne, Michael Lynskey. (11/05/2020). Assessing the public health impacts of legalizing recreational cannabis use: the US experience. World Psychiatry. https://doi.org/10.1002/wps.20735
[26] Gingiss, Phyllis M. et al. (23/05/2006). Bridge-It: A System for Predicting Implementation Fidelity for School-Based Tobacco Prevention Programs. Prevention Science, 197-207. https://doi.org/10.1007/s11121-006-0038-1

[27] Vigna-Taglianti, Federcia D. et al. (2014). "Unplugged," a European school-based program for substance use prevention among adolescents: overview of results from the EU-Dap trial. New Dir Youth. Dev. https://pubmed.ncbi.nlm.nih.gov/24753279/
[28] Alarco-Roslaes, Raquel et al. (2021). Effects of a School-Based Intervention for Preventing Substance Use among Adolescents at Risk of Academic Failure: A Pilot Study of the Reasoning and Rehabilitation V2 Program. Healthcare. https://doi.org/10.3390/healthcare9111488
[29] Ransing, Ramdas et al. (2022). Current state of cannabis use, policies, and research across sixteen countries: cross-country comparisons and international perspectives. Trends in Psychiatry and Psychotherapy. https://doi.org/10.47626/2237-6089-2021-0263

The Science Behind the Trends: Evaluating Popular Acne Remedies on Social Media Against Clinical Evidence By Rasil Sheikh

Abstract

Acne vulgaris is one of the most prevalent dermatological conditions, for approximately 50 million Americans each year, up to 85% of whom are adolescents. Beyond these physical symptoms, acne imposes significant psychosocial burdens, including scarring, anxiety, and depression. Meanwhile, adolescents are among the most active users of social media, with millions engaging with platforms like TikTok and Instagram. Increasingly, they turn to these platforms, not dermatologists, for beauty advice. This has fueled the viral spread of "DIY" acne remedies, which range from citrus-based masks to detox drinks. These are often promoted by uncredentialed influencers. To evaluate the accuracy and safety of these widely circulated remedies, we conducted a two-part investigation: (1) a content analysis of the most-viewed acne treatments on TikTok, Instagram, and YouTube, and (2) a systematic literature review using PubMed and Google Scholar to assess the scientific evidence for each treatment's safety and efficacy. Our findings reveal a clear contrast: the vast majority of popular remedies (such as lemon juice, "skin detox" recipes) lack clinical support and, in many cases, pose risks such as chemical burns or irritation. In contrast, evidence-based treatments such as benzoyl peroxide, topical retinoids, and salicylic acid are consistently underrepresented in social media content despite their proven effectiveness. These results highlight a major gap between dermatological science and social media narratives, with real implications for adolescent skin health. Targeted digital literacy and public health outreach may be necessary to bridge this divide.

Keywords

Acne vulgaris, TikTok skincare trends, Instagram beauty remedies, Remedies, Social Media

Introduction

Acne vulgaris is one of the most prevalent dermatological conditions globally, particularly affecting adolescents. In the United States alone, approximately 50 million people experience acne annually, with up to 85% of teenagers and young adults developing lesions at some point [1]. Beyond its physical manifestations, acne can lead to lasting scarring and profound psychosocial consequences, including anxiety, depression, low self-esteem, and even suicidality, which marks it as a significant public health concern [1].

At the same time, today's adolescents are immersed in digital culture. According to Pew Research, 90% of U.S. teens use YouTube, and over 60% engage regularly with TikTok, Instagram, or Snapchat [2]. Again and again, these platforms serve as their primary source for skincare advice, often in place of medical professionals [3]. While social media offers easy access to information, it also provides a fertile ground for misinformation. Dermatologists have raised concerns about the surge of pseudoscientific content circulating online where acne "cures"

frequently emphasize quick fixes, dietary myths, and unproven home remedies over evidence-based treatments [4,5,6].

Recent studies also confirm this trend: viral content that includes hashtags like #acne or #clearskin commonly promotes DIY approaches such as lemon juice applications or detox drinks [7]. These are often endorsed by influencers without medical credentials [8]. These videos garner millions of views, shaping skincare beliefs and behaviors in ways that are potentially harmful and ineffective [9,10,11].

Literature Review

Our systematic literature review assessed the scientific evidence supporting various acne remedies popularized on social media. The results are summarized in the table below. For Lion Pair Cream, no peer-reviewed clinical trials or systematic reviews were found, categorizing it as having no scientific support, with safety risks unknown beyond general topical cream concerns [12,13]. Tretinoin demonstrated robust evidence from numerous randomized controlled trials (RCTs), confirming its efficacy in reducing both inflammatory and non-inflammatory acne lesions [13]. It is widely recommended in clinical guidelines, though common adverse effects include dryness, peeling, redness, and increased photosensitivity. Benzoyl peroxide and salicylic acid both received strong support from meta-analyses and Phase III clinical trials, confirming their efficacy and safety profiles [14,15]. Benzoyl peroxide has been further validated by recent studies showing no association with cancer risks, although potential safety issues include skin irritation, dryness, fabric bleaching, and rare benzene contamination [16,17]. Salicylic acid, similarly effective, is often used in combination with benzoyl peroxide [18]. PanOxyl, a brand containing benzoyl peroxide, inherits strong scientific backing due to its active ingredient and is frequently recommended by dermatologists despite limited brand-specific studies [19]. In contrast, detox waters and juices, notably those with lemon, lack clinical trial evidence supporting their effectiveness against acne vulgaris. Dermatological sources highlight safety concerns including chemical burns, skin irritation, and increased photosensitivity, categorizing these remedies as unsupported scientifically [20]. Turmeric and honey masks received moderate preliminary support based on small pilot studies and systematic reviews, indicating potential antimicrobial and anti-inflammatory properties [21][22]. However, larger human trials are still needed. Safety issues noted include allergic reactions, skin staining, and variability in dosage and outcomes [22]. Murad Acne Cream lacks independent randomized controlled trials, but its key ingredient, salicylic acid, is supported by moderate evidence. Company-conducted short-term studies suggest a high level of user satisfaction, yet standard safety concerns related to salicylic acid such as dryness and skin irritation apply [23, 24].

Table 1. Relationship between the Social media remedy and scientific evidence for that remedy.

Remedy	**Evidence Strength**	**Safety Notes**
Lion Pair Cream	No Support	Unknown beyond cream concerns
Tretinoin	Strong Support	Dryness, irritation, photosensitivity
Benzoyl Peroxide+ Salicyclic Acid	Strong Support	Irritation, drying, bleaching, rare benzene, contamination.
Panoxyl	Strong Support	Same as Benzoyl peroxide
Detox water/lemon juice	No Support	Chemical burns, photosensitivity, irritation
Turmeric/Honey Mask	Moderate Support	Allergic Reactions, staining
Murad Acne Cream	Moderate Support	Dryness, irritation(salicylic acid effects)

Adapted and created by the authors (Data from Tiktok & Instagram).

Methodology

Social Media Trend Analysis

We conducted a structured content analysis of posts from TikTok, YouTube, Instagram and Reddit. Searches included acne-related keywords (e.g. “acne cure,” “pimple hack,” “clear skin hack”) and popular hashtags (#acne, #clearskin, etc.) which follow methods similar to prior work. We identified the top engagement posts (50-100 per platform) and manually coded each for the key remedy or product featured, influencer type, view/like counts, and claims made (e.g. “X cured my acne overnight”). From this, we built a spreadsheet of the 30-40 most viral remedies, and included the top 7 most viral ones in a table and their purported benefits/safety claims (see excerpt below).

Scientific Literature Review

For each identified remedy, we searched PubMed/Google Scholar using “[remedy] AND acne vulgaris.” This included specific terms like “tea tree oil,” “turmeric AND acne,” “benzoyl peroxide AND acne,” etc. We prioritized systematic reviews, Cochrane analyses, and RCTs. We extracted data on study design, efficacy outcomes, and adverse effects. Remedies were then classified by the strength of evidence (e.g. “strong support,” “moderate support,” “no support”) and known safety issues.

Data Synthesis

We summarized social-media popularity (views/likes) and scientific support side by side. Numerical metrics from platforms (e.g. TikTok views) were recorded from the content analysis. For evidence grading, we relied on cited trials/guidelines, noting where no trials exist. Safety summaries drew on clinical reports and dermatology resources.

Results

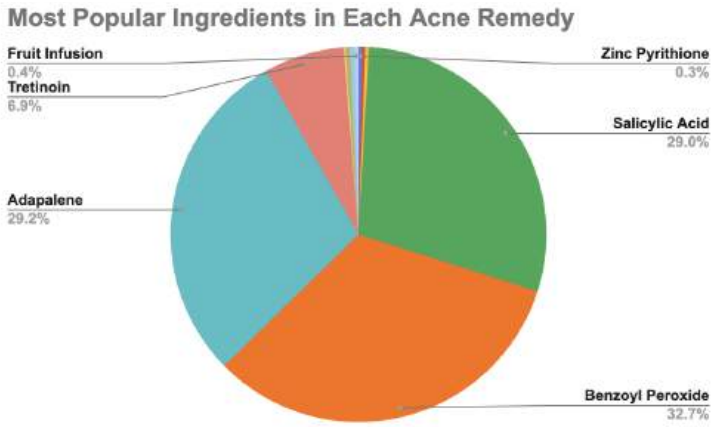
The table below compares popular viral acne remedies with scientific evidence:

Table 2. Correlation between Popular Viral Acne Remedies and the Scientific Evidence

Remedy	**Social Media Trend (Approx)**	**Claimed Effects**	**Scientific Evidence**	**Safety Concerns**
Lion Pair Cream	308K+ likes on TikTok	Acne is gone; Acne scars are gone	Limited or mixed evidence; varies by ingredient	Unknown beyond cream concerns; one should seek a dermatologist before using it
Tretinoin	417K+ likes on tiktok	Reduces inflammation and wrinkles	Strong clinical support. Prescription retinoid that increases cell turnover and reduces acne	High potential for dryness, photosensitivity, and irritation.
Benzoyl peroxide, and Salicylic Acid	546K+ likes on Tiktok	Reduces acne	Strong support. FDA approved acne treatment shown to reduce acne causing bacteria and inflammation.	Can cause dryness, irritation, and bleaching of fabrics.
Panoxyl	211K+ likes on Tiktok	Skin Glows	Contains Benzoyl Peroxide which is strong evidence for acne reduction.	May cause irritation or dryness to the eyes; avoid contact with eyes.
“Detox” water / juices (e.g. lemon water)	22,170 likes on Instagram	“Detoxifies skin internally, prevents acne”	No evidence. Drinking water or juices has no known effect on acne clearance. These claims stem from general health myths. There are no trials testing “detox” beverages for acne.	No evidence. Drinking water or juices has no known effect on acne clearance. These claims stem from general health myths. There are no trials testing “detox” beverages for acne. Harmless if non-citrus, but citrus juices still risk skin irritation if applied topically. Ingesting excessive citrus can upset the stomach.
Turmeric / honey mask or patch	10,639 likes on Instagram	“Anti-inflammatory; reduces redness/bumps”	Limited preliminary evidence. Curcumin (turmeric’s active) has shown antimicrobial and anti-inflammatory effects in lab and animal studies. Some small trials hint it may help inflammatory acne, but no robust human RCT confirms efficacy	Generally safe to eat; topical turmeric can stain skin and sometimes cause contact dermatitis. Honey may occasionally irritate acne-prone skin. High-dose supplements can cause GI upset.
Murad Acne Cream	2K+ likes	Makes acne bumps flatter	Limited evidence.	May cause dryness to the skin. Seek a dermatologist.

Adapted and created by the authors (Data from Tiktok & Instagram)

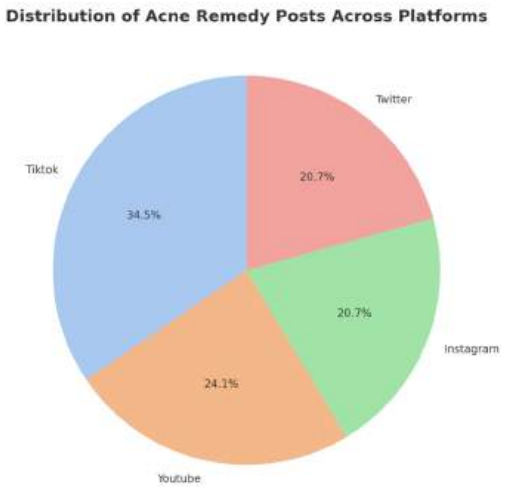
Figure 1. Most Popular Ingredients in Each Acne Remedy



Adapted and created by the authors (Data from Tiktok & Instagram).

The pie chart above further reveals that the majority of viral remedies in our dataset feature dermatologist-approved ingredients. Benzoyl peroxide appears most frequently (32.7%), followed closely by adapalene (29.2%) and salicylic acid (29.0%). These ingredients are widely supported by clinical research and are recommended in most acne treatment guidelines, indicating that while social media content often lacks context or professional oversight, many users still gravitate toward products with scientifically validated ingredients. In contrast, ingredients like tretinoin (6.9%),despite strong evidence, appear less frequently, possibly due to their prescription-only status. Meanwhile, ingredients such as zinc pyrithione (0.3%) and fruit infusions (0.4%) are rare and reflect more experimental or anecdotal approaches.

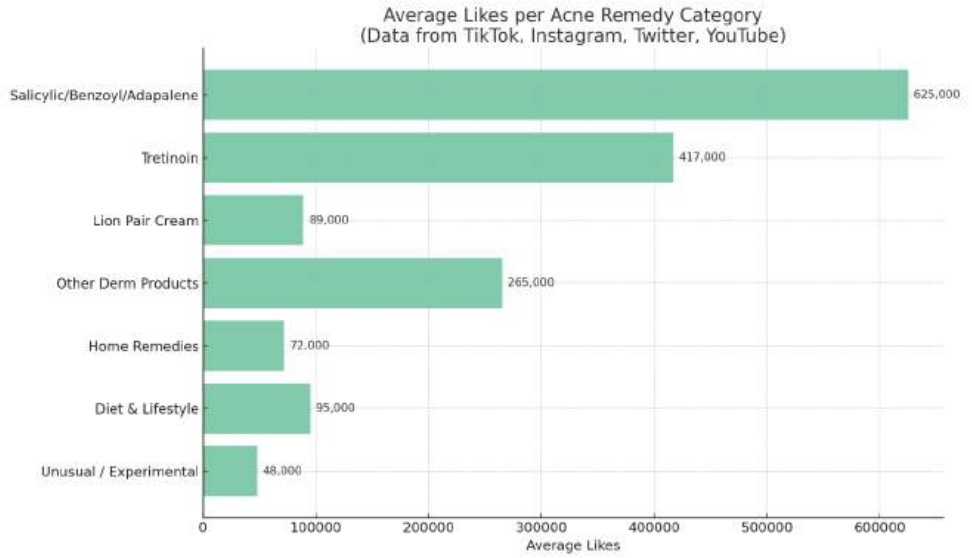
Figure 2. Distribution of Acne Remedy Posts Across Platforms



Adapted and created by the authors (Data from Tiktok & Instagram).

As seen in the Pie Chart in Figure 2, TikTok accounts for the largest share of viral acne remedy posts (34.5%), followed by YouTube (24.1%), while Instagram and Twitter each contribute 20.7%. This distribution reflects how platforms known for short-form or viral content, especially TikTok, play an outsized role in shaping skincare narratives online. Our research confirms that many TikTok posts promote trendy but unverified remedies, such as Lion Pair Cream or DIY acne routines, often accompanied by dramatic before-and-after claims. YouTube, while more balanced, still hosts a mix of credible dermatology advice and anecdotal or influencer-driven solutions. Instagram posts frequently center on holistic or aesthetic remedies like turmeric masks or detox waters, which tend to lack strong clinical backing. Twitter posts, being shorter and less detailed, often spread simplified advice without scientific context. The chart reinforces the key concern in our research of how the most popular platforms for acne remedy content are also those where scientific evidence is often weakest, thus why misinformation can spread rapidly.

Figure 3. Average Likes Per Acne Remedy Category



Adapted and created by the authors (Data from Tiktok & Instagram).

This bar graph shows the average number of likes each acne remedy category received across TikTok, Instagram, Twitter, and YouTube. Clinically proven treatments like salicylic acid, benzoyl peroxide, and adapalene had the highest engagement (~625K likes), followed by tretinoin and dermatologist-recommended products such as PanOxyl and CeraVe. Interestingly, diet and lifestyle changes and home remedies (such as turmeric, lemon) also gained considerable attention, despite weaker scientific support. Lion Pair Cream, a popular Japanese product, received moderate interest. Meanwhile, experimental methods like anti-dandruff shampoo or topical insulin had lower, but still notable, engagement. Overall, the data suggests that social media users are drawn to both science-backed treatments and trending natural or unconventional

solutions. Our analysis illustrates a clear contrast between the acne remedies most widely circulated on social media and those supported by dermatological research. Across platforms (particularly TikTok and Instagram) we found that the majority of viral acne treatments lacked strong scientific validation. Although medically endorsed treatments such as topical retinoids (e.g., tretinoin), benzoyl peroxide, and salicylic acid were present in the dataset, they were relatively underrepresented and often embedded within broader, less substantiated skincare routines. For example, a post might mention salicylic acid, but center the routine on more attention grabbing and non-clinical elements like ice facials, turmeric-honey masks, or "detox" drinks. Among the most frequently viewed remedies were Lion Pair Cream, tretinoin, salicylic acid, lemon face mask, and panoxyl. Many of these, particularly Lion Pair Cream and lemon based treatments, were praised for immediate results, yet lacked strong scientific and clinical evidence and posed safety concerns such as skin irritation and acid burns. In contrast, dermatologist-approved options like tretinoin and salicylic acid, while backed by strong scientific support, often received less engagement. Based on this disparity, we were able to conclude that social media algorithms and influencer narratives prioritize novelty and virality over evidence-based care, and that the popularity of a certain remedy does not correlate with its medical credibility, this case being acne treatment. The quantitative analysis of our dataset suggests that approximately 70% of the top viewed remedies lacked substantial clinical support.

Furthermore, roughly 35-40% posed potential harm, particularly when used incorrectly or in combination with other harsh substances. Some documented risks included chemical burns from lemon juice, excessive dryness or peeling from retinoids, and contact dermatitis from turmeric-based masks. In rare but concerning cases, we found anecdotal evidence of users experiencing worsened breakouts, long-term skin discoloration, or allergic reactions after following influencer-endorsed routines without professional guidance. The discrepancy between popularity and scientific reliability raises serious public health concerns.

Discussion

Adolescents and young adults who are already vulnerable to the psychosocial impacts of acne may be misled by these viral trends, delaying or entirely avoiding treatments that are both safe and effective. In some cases this misinformation could lead to worse outcomes, including prolonged acne flares, increased risk of scarring, or new injuries from inappropriate topical applications. Moreover, the normalization of untested remedies on social media contributes to widespread confusion about how acne actually develops and what treatments are supported by medical science. Dermatologists have warned that the credibility of health content on social platforms varies widely depending on the source. Posts that are created by board certified dermatologists or other licensed professionals often provide accurate and balanced information, yet these are outnumbered by content from influencers with no formal training.

Our findings support these concerns: in our dataset, posts from non-medical creators were not only more common, but also more likely to include exaggerated claims such as “clears your skin overnight” or “cures acne naturally in days”. Multiple peer reviewed studies on health

misinformation online confirm that adolescents are especially susceptible to such claims, particularly when they are visually appealing and endorsed by attractive influencers. This points to a broader issue in digital health literacy. While social media has the potential to democratize health education, it also risks amplifying pseudoscience when left unmoderated. Our results reinforce the need for stronger evidence based communication on platforms heavily used by youth. Some interventions such as promoting verified dermatologists on TikTok or including disclaimers on unverified medical claims may help reduce harm. Narrowing the gap between dermatological science and social media skincare culture will require collaboration between medical professionals, educators, public health organizations, and the platforms themselves.

Acknowledgements

We would like to thank our parents for supporting us while writing this.

Limitations

Our study has several important limitations. Firstly, the content analysis focused only on the most popular posts across selected social media platforms, which may not fully represent the breadth of available content or capture less visible but potentially influential posts. Secondly, popularity metrics (views, likes) were more likely than not influenced by factors unrelated to the perceived effectiveness or safety of remedies, such as influencer popularity or platform algorithms, potentially skewing the data. Thirdly, the literature review was limited to articles available in PubMed and Google Scholar databases, possibly overlooking relevant studies published in non-indexed journals or in languages other than English.

Conclusion

Our analysis reveals a significant difference between the most viral acne remedies on social media and those supported by clinical dermatology. Approximately 70% of the top viewed treatments lacked substantial scientific backing, and over one third posed potential safety risks such as skin irritation, acid burns, or allergic reactions, often when misused or combined with other harsh ingredients. Remedies such as lemon face masks and Lion Pair Cream were widely promoted despite little to no peer-reviewed evidence. Meanwhile clinically endorsed treatments such as benzoyl peroxide, carbon facials, and salicylic acid, while highly effective, were relatively underrepresented and often overshadowed by more sensational and influencer-driven routines.

This gap sparks the urgent need for digital health literacy and intervention: social media platforms could prioritize vetted content, dermatologists might expand their online presence, and educators and families can help adolescents critically evaluate skincare claims. Increasingly, teens should be equipped with accurate, evidence based guidance to protect their skin health and reduce the harm caused by viral misinformation.

Works Cited

American Academy of Dermatology. (n.d.). Acne: Diagnosis and treatment. https://www.aad.org/public/diseases/acne/derm-treat/treat

American Academy of Dermatology. (n.d.). Skin conditions by the numbers. https://www.aad.org/media/stats-numbers

Atkin, N. J. (2020). From high-flying trapeze to highly infectious disease: The history of dermatology in the circus. Clinical Dermatology, 38(5), 654-660. https://doi.org/10.1016/j.clindermatol.2020.09.006

Adebesin, F., Smuts, H., Mawela, T., Maramba, G., & Hattingh, M. (2023). The role of social media in health misinformation and disinformation during the COVID-19 pandemic: Bibliometric analysis. JMIR Infodemiology, 3(1), Article e48620. https://doi.org/10.2196/48620

Bal, Z. I., Karaosmanoglu, N., Temel, B., Bas, V. M., Karali, M. G., Ergun, E. Z., Avci, E. B., Tellal, E. S., Yuksel, E. I., & Nalbant, S. (2025). Trust in dermatologists versus social media influencers among acne patients. Cureus. https://doi.org/10.7759/cureus.57021

Dréno, B., Bettoli, V., Araviiskaia, E., Ochsendorf, F., & Layton, A. (2018). The influence of exposome on acne. Journal of the European Academy of Dermatology and Venereology, 32(9), 1445-1455. https://doi.org/10.1111/jdv.14820

Ertekin, S. S., Salici, N. S., Manav Bas, V., Karali, M. G., Ergün, E. Z., Avcı, E. B., Tellal, E. S., & Yüksel, E. I. (2024, July). Influence of social media and internet on treatment decisions in adult female acne patients: A cross-sectional survey study. Dermatology Practical & Conceptual, 14(3), Article e2024156. https://doi.org/10.5826/dpc.1403a156

George Washington University. (2023). Study finds TikTok filled with acne misinformation. https://mediarelations.gwu.edu/study-finds-tiktok-filled-acne-misinformation

Hughes, M. M., Shaw, K. A., Patrick, M. E., & DiRienzo, M. (2023). Adolescents with autism spectrum disorder: Diagnostic patterns, co-occurring conditions, and transition planning. Pediatrics, 152(3), Article e2022060085. https://doi.org/10.1542/peds.2022-060085

Hunt, I. de V., & Linos, E. (2022). Social media for public health: Framework for social media-based public health campaigns. Journal of Medical Internet Research, 24(12), Article e42179. https://doi.org/10.2196/42179

Leyden, J., Stein-Gold, L., & Weiss, J. (2017). Why topical retinoids are mainstay of therapy for acne. Dermatology and Therapy, 7(3), 293-304. https://doi.org/10.1007/s13555-017-0185-2

Liu, Y., Dan, Y., Yang, J., He, X., Liu, J., Yi, Y., Chen, X., Yin, X., Song, W., Niu, Y., Zheng, Y., & Ai, Y. (2025). Clinical efficacy of a salicylic acid–containing gel on acne management and skin barrier function: A 21‐day prospective study. Journal of Cosmetic Dermatology, 24(7), Article e70353. https://doi.org/10.1111/jocd.70353

Medical News Today. (n.d.). Turmeric face masks. https://www.medicalnewstoday.com/articles/turmeric-face-mask

Medical News Today. (2024, June 10). Everything you need to know about cystic acne. https://www.medicalnewstoday.com/articles/103258#prevention

Montgomery, K. C., & Chester, J. (2001). Youth and digital media: A policy research agenda. Journal of Broadcasting & Electronic Media, 45(1), 138-150. https://doi.org/10.1016/s1054-139x(00)00130-0

O'Connor, C., O'Grady, C., & Murphy, M. (2024). Spotting fake news: A qualitative review of misinformation and conspiracy theories in acne vulgaris. Clinical and Experimental Dermatology, 49(1), 1-8. https://doi.org/10.1111/ced.15222

Ogwu, M. C., & Izah, S. C. (2025, March 1). Honey as a natural antimicrobial. Antibiotics, 14(3), Article 255. https://doi.org/10.3390/antibiotics14030255

Pappas, A. (2009). The relationship of diet and acne. Dermato-Endocrinology, 1(5), 262-267. https://doi.org/10.4161/derm.1.5.10192

Pew Research Center. (2023). Teens, social media and technology 2023. https://www.pewresearch.org/internet/fact-sheet/teens-and-social-media-fact-sheet/

The Dermatology Digest. (n.d.). PanOxyl product recommendations. https://thedermdigest.com/survey-panoxyl-named-no-1-dermatologist-recommended-otc-benzoyl-peroxide-brand/

Wojtara, M. S. (2023). Use of social media for patient education in dermatology: Narrative review. JMIR Dermatology, 6, Article e42609. https://doi.org/10.2196/42609

Yousaf, A., Hagen, R., Delaney, E., et al. (2021). The influence of social media on acne treatment: A cross-sectional survey. Journal of the American Academy of Dermatology. https://doi.org/10.1111/pde.14091

Zaenglein, A. L., Pathy, A. L., Schlosser, B. J., Alikhan, A., Baldwin, H. E., Berson, D. S., Bowe, W. P., Chien, A. L., Cohen, J. L., Collyer, A. J., Del Rosso, J. Q., Draelos, Z. D., Gold, L. F., Granger, C., Harper, J. C., Keri, J. E., Kircik, L. H., Levy, M. L., Leyden, J. J.,... Bhushan, R. (2016). Guidelines of care for the management of acne vulgaris. Journal of the American Academy of Dermatology, 74(5), 945-973. https://doi.org/10.1016/j.jaad.2006.08.048

Zhang, Z., & Zhou, M. (2024, August 1). The impact of social media information exposure on appearance anxiety in young acne patients: A moderated chain mediation model. Frontiers in Psychology, 15, Article 1409980. https://doi.org/10.3389/fpsyg.2024.1409980

Sanitation Reformers in 19th-century Britain: Pioneers of Environmental Justice and Public Health By Anya Kothari

Abstract

Despite significant global progress in sanitation, there is limited attention on the historical roots of sanitary science and its foundational role in shaping modern public health in Europe and globally. This paper examines the role of 19th-century British sanitation reformers, such as Edwin Chadwick and John Snow, who responded to the public health crises brought on by rapid industrialization and urbanization. The sanitation reformers were pioneers who explored uncharted territory. Drawing on evidence from personal correspondence, research, speeches, government reports, and other primary and secondary sources, this study deepens understanding of how the reformers linked environmental conditions to disease, reframed poverty as a systemic issue rather than a moral failing, and positioned sanitation as a matter of social justice and state responsibility. Although the sanitation reformers evolved through many theories, often making mistakes, their work fostered advancements in legislation, infrastructure, and public awareness. The reformers also operated against a backdrop of class struggle, which influenced the politics of sanitation reform. However, their overall objective remained unchanged–to save lives by promoting equality and government oversight of public health. Although it was challenging to trace the different viewpoints at the time and figure out the sequence of events that catalyzed reform, extensive research led to the identification of useful primary sources for analysis and the crafting of a cohesive argument. The work of the sanitation reformers is an important topic to study as it catalyzed the field of public health and medical advancements. By integrating science, politics, and economic policy, the sanitation reformers not only redefined the relationship between class and health but also laid the groundwork for the British welfare state and modern public health systems worldwide.

"A dirtier or more wretched place he had never seen. The street was very narrow and muddy, and the air was impregnated with filthy odours."[3]

Oliver on his journey to London in Charles Dickens' *Oliver Twist*

Introduction

Most people in the developed world live in areas with reliable sewage removal and widespread access to clean water. Citizens often take these fundamental human rights for granted. However, the people of 19th-century Great Britain experienced a sanitation crisis that shook their society and disproportionately affected the poor. Prompted by the Industrial Revolution, rapid urbanization led to unhygienic environments, disease, and poverty associated with filth and squalor. Social reformers such as Edwin Chadwick and John Snow, among others, investigated the conditions of the poor and working class, exploring the relationship between the environment and health. Recognizing the need for sanitation to promote economic growth, maintain social order, and support the working class, Chadwick published the *Report on the Sanitary Condition of the Labouring Population of Great Britain*. Through his work, Chadwick initiated a movement and showed that unsanitary conditions caused destitution, not the other way around. The poor were, in fact, not inherently immoral or filthy, as was the prevailing opinion, and improving public health would benefit society overall. Although sanitation reformers were not always correct, sometimes operated from a narrow class-based perspective, and failed to implement universal clean water systems and sewage disposal immediately, they were not afraid to experiment–advancing knowledge, legislation, and awareness. By bringing together politics, economics, and social issues, the reformers generated sanitation reform and prosperity, proving that social class should not determine health outcomes. The reformers' evidence-driven ideas altered society's view of the poor, linked disease to unsanitary environments, and established the notion of interconnected public health with a central government as the responsible arbiter. As a result, the reformers laid the foundation for the modern British welfare state and catalyzed public health movements, transforming disease prevention, social justice, and life expectancy globally.

Britain in the 19th Century and the Industrial Revolution

The Industrial Revolution ushered in rapid urbanization, altering the way of life in Great Britain. The rise of factories and industries created jobs and demand for labor, drawing people from rural areas to cities. As a result, the percentage of British people living in cities rose dramatically during the first half of the nineteenth century. At the end of the 18th century, when industrialization began, approximately 30 percent of the British population lived in urban areas. By 1841, that number had increased to 50 percent.[4] London was the only city in Britain with a population over 100,000 in 1801. However, by 1851, ten towns in the country also had over

[3] Charles Dickens, *Oliver Twist* (Global Publishers, 2023), 73, Kindle.

[4] Ian Morley, "City Chaos, Contagion, Chadwick, and Social Justice," *The Yale Journal of Biology and Medicine* vol. 80, 2 (2007): 61, https://pmc.ncbi.nlm.nih.gov/articles/PMC2140185/.

100,000 residents.[5] The dramatic increase in urban populations was problematic, exceeding housing capacity and leading to health issues and diseases such as cholera. Death rates increased, and life expectancies declined. Although a limited number of houses had piped water, most relied on a shared tap or standpipe for their water supply. Human waste overflowed into the streets and was discharged into rivers from drains. There was limited garbage collection, and filth was everywhere.[6] Michael Faraday, a physicist, called the Thames River in London "an opaque brown fluid" and a "fermenting sewer."[7] The Thames was also satirized by artists. Filthy conditions prevailed in small towns and villages beyond large cities, such as London, Manchester, and Glasgow.

Although inadequate sanitation affected everyone, the poor suffered disproportionately. The mortality rate from cholera was 25 times higher in the working-class districts of Rotherhithe than in the upmarket area of Hampstead.[8] To address the plight of the poor, the Poor Laws, created in 1601, remained in effect in the early 1800s. The laws were designed to help orphans, the elderly, the sick, and the unemployed. However, the Poor Laws fostered public discontent because the government lacked sufficient funds to provide adequate aid. Edwin Chadwick, a lawyer, was asked by the government commission, the Royal Commission for Inquiring into the Administration and Practical Operation of the Poor Laws, to modify the outdated legislation. Previously, Chadwick was an assistant to the utilitarian philosopher Jeremy Bentham. Bentham believed "the greatest happiness of the greatest number…is the measure of right and wrong," and Chadwick, greatly influenced by Bentham, adopted this thinking.[9] Chadwick looked for opportunities that had the greatest impact on the largest number of citizens. Furthermore, for Chadwick, "greatest national profit" was the economic equivalent of the "greatest happiness principle."[10] After joining the Royal Commission to inquire into the Poor Laws, Chadwick visited more than 3,000 out of 15,000 parishes to gather information and approach the problem methodically. Believing that the existing system of Poor Laws in England was "destructive of industry, forethought and honesty of the labourers" and destructive to "the mutual goodwill and happiness of all" since it "collects and chains down the labourers in masses," Chadwick became increasingly concerned "these evils are becoming more overwhelming in their magnitude and less susceptible of cure."[11] As a result, Chadwick took decisive action. Ultimately, he became secretary to the Poor Law Commission and executed the 1834 Parliament-enacted Poor Law Amendment Act, whereby the able-bodied poor did not receive payment relief; instead, they

---

[5] Anthony S. Wohl, *Endangered Lives: Public Health in Victorian Britain* (Cambridge: Harvard University Press, 1983), 3.
[6] Fred C. Pampel, *The Struggle for Public Health* (Baltimore: Johns Hopkins University Press, 2024), 10.
[7] Pampel, *The Struggle for Public Health*, 10.
[8] Paul Bingham, Neville Q. Verlander, and M. J. Cheal, "John Snow, William Farr and the 1849 Outbreak of Cholera that Affected London: a Reworking of the Data Highlights the Importance of the Water Supply," Public Health 118, no. 6 (2004): 387. https://doi.org/10.1016/j.puhe.2004.05.007.
[9] Pampel, *The Struggle for Public Health*, 16.
[10] Samuel Edward Finer, *The Life and Times of Sir Edwin Chadwick (*London: Methuen & Co. LTD., 1952), 23.
[11] Frederick Purdy, "The Statistics of the English Poor Rate before and since the Passing of the Poor Law Amendment Act," *Journal of the Statistical Society of London* 23, no. 3 (1860): 305. https://doi.org/10.2307/2338499.

were required to live in "workhouses." The strict workhouses were meant to make the poor more productive and to discourage them from relying on aid, which Chadwick believed was better for everyone. Chadwick asserted that the workhouse was an applicable solution for most cases. The workhouse served as a hospital, an asylum, and as a solution to "ordinary destitution."[12] The workhouses were so harsh and unappealing that paupers would often decide not to work there. Those who continued to receive aid were required to reside in the workhouses, where children were separated from their parents, and meager food was served. The conditions were dire, and paupers were often "placed together wet, filthy, infested with vermin, and diseased, in a small, unventilated space."[13] Chadwick's extreme measures made the poor unhappy, and the rich did not like his intervention in government authority. *The Times of London* dubbed the amended Poor Law Act "the starvation act."[14] Other reformers, such as William Farr, a statistician and epidemiologist, maintained that the workhouses were destructive to the health of the poor, and he highlighted the problem of starvation as a cause of death and disease. Farr honed in on 63 deaths from starvation and concluded that the workhouse diet only provided three-quarters of the nutritional requirements for an individual.[15] Farr and Chadwick debated in depth about the relationship between starvation and disease. Farr used a broader definition of starvation and said, "few die from the absolute want of food…many die, or drag on a miserable existence upon insufficient, innutritious diet."[16] For example, he noted the lack of warmth as a contributing factor to the deaths. Because Farr disagreed with Chadwick, Chadwick undertook a deeper and more thorough exploration of the factors that contributed to the diseases, as well as a more comprehensive evaluation of diet and working conditions.

However, because Chadwick believed that large enterprises were needed to advance the British economy, he continued to advocate for strict conditions and the continuation of child labor through the Factory Act. Chadwick's utilitarian belief in a free market economy, coupled with his evidence-based inquiry, underpinned his actions. In effect, the new Poor Law was "the first great piece of legislation based upon scientific or economic principles."[17] Chadwick's work on the Poor Laws aimed to demonstrate that the poor were not inherently different from others and could be rehabilitated and utilized to contribute to the betterment of the economy and society for the greater good.

Sanitation Reform Takes Root

---

[12] Edwin Chadwick, "Poor Law Administration, Its Chief Principles and Their Results in England and Ireland as Compared with Scotland." *Journal of the Statistical Society of London* 27, no. 4 (1864): 494. https://doi.org/10.2307/2338611.

[13] Henry Mayhew, *London Labour and the London Poor Volume 3* (London: Griffin, Bohn & Company, 1861), 367.

[14] Christopher Hamlin, "Could You Starve to Death in England in 1839? The Chadwick-Farr Controversy and the Loss of the "Social" in Public Health," *American Journal of Public Health* 30, vol. 85, 6 (1995): 857. doi:10.2105/ajph.85.6.856.

[15] Hamlin, "Could You Starve to Death in England in 1839? The Chadwick-Farr Controversy and the Loss of the "Social" in Public Health," 858.

[16] Hamlin, "Could You Starve to Death in England in 1839? The Chadwick-Farr Controversy and the Loss of the "Social" in Public Health," 858.

[17] Finer, *The Life and Times of Sir Edwin Chadwick*, 69.

Guided by the principle that improved public health could also lead to greater prosperity and social order, and confronted with growing and widespread filth, Chadwick shifted his attention away from workhouses and child labor to the larger issue of sanitation. He began collaborating with a fellow follower of Bentham, Thomas Southwood Smith. Since Chadwick did not have a background in science or medicine, he turned to others, such as Southwood Smith for, medical insights. Southwood Smith was a physician who worked at the London Fever Hospital in the impoverished East End of London. Southwood Smith observed the workers' dismal living conditions, medical afflictions, and poor health there. He concluded that the filth in which they lived had a detrimental effect on their physical health. The poor in areas such as East London were "dying at an average of 29 years less than the wealthy."[18] Espousing the widely accepted idea that epidemic diseases came from airborne miasmas, Southwood Smith developed the "filth theory of disease," which asserted that the origin of the miasma was filth in a specific neighborhood or community. In his 1830 book, *Treatise on Fever,* he concluded that the decomposition of vegetable and animal matter gave off a new compound that produced fever and disease.[19] In his 1838 report "On Some of the Physical Causes of Sickness and Mortality to which the Poor are Particularly Exposed," he showed "while in the districts of London occupied by the wealthier classes streets had been widened, drainage improved, and nuisances removed, nothing whatever had been done to effect similar remedies in the poorer districts."[20] The poor had few or no means to avoid filth and little or no choice of their homes.[21] The differences in the living conditions among the classes were stark. Deeply concerned, Chadwick visited the East End of London to investigate and gather information about the poorer sections of the city. Chadwick contracted a fever from typhus and almost died, his own personal experience reinforcing his insight that filth caused disease.[22] His own observations and Southwood Smith's work inspired Chadwick to document the link between poverty, poor environments, and the spread of diseases, ultimately leading to reform.

Observing that unsanitary conditions caused sickness and mortality, Chadwick realized that the stigma of filth that the poor carried was wrong. He saw the terror that cholera incited in Britain after its destructive spread and how it did not respect social class boundaries.[23] The cholera epidemic of 1831-1832 showed the urgency of implementing sanitary reform in both "town and country," given the "grave and epidemic" nature of the disease.[24] Thus, Chadwick turned to understanding and preventing diseases caused by environmental factors. The Poor Law Board sent a letter to Lord John Russell, the Home Secretary, in 1838 saying that changing poor

---

[18] Morley, "City Chaos, Contagion, Chadwick, and Social Justice," 69.
[19] Thomas Southwood Smith, *A Treatise on Fever* (London: Longman, Rees, Orme, Brown, and Green, 1830), 349, https://archive.org/details/b33093519_0002/page/348/mode/2up.
[20] Ashworth Underwood, "The Centenary of British Public Health: Rise of Health Legislation in England and in London," *The British Medical Journal* 1, no. 4557 (1949): 891, accessed May 1, 2025, https://doi.org/10.1136/bmj.1.4557.890.
[21] Underwood, "The Centenary of British Public Health: Rise of Health Legislation in England and in London," 890.
[22] Pampel, *The Struggle for Public Health*, 29.
[23] Morley, "City Chaos, Contagion, Chadwick, and Social Justice," 67.
[24] Underwood, "The Centenary of British Public Health: Rise of Health Legislation in England and in London," 890.

rates meant focusing on issues "caused by nuisances by which contagion is generated and persons are reduced to destitution."[25] In 1838, Chadwick commissioned medical practitioners James Kay-Shuttleworth and Neil Arnott to investigate how sanitary measures could prevent fever in London.[26] Back to his utilitarian principles, Chadwick promoted the idea that prevention was superior to curing.[27] William Farr collaborated with Chadwick to analyze mortality data and use statistics to solidify further the link between sanitation and public health. Chadwick argued that sanitation was a solution to poverty, as well as to a lack of sewer systems, poor ventilation, and overcrowding, which caused disease. Disease then led to poverty and more filth, which made people even sicker, perpetuating a vicious cycle of poverty.

Sanitary Report - A Novel Approach

As he became increasingly convinced that sanitation was a problem in British society, Chadwick built upon the collective sanitation work that had been done, focusing on documentation, statistics, and the dissemination of information. In 1839, Queen Victoria instructed Chadwick to investigate disease among the working classes in parts of England and Wales. Chadwick utilized statistical evidence to highlight disparities in life expectancy based on class and residence. His 1842 *Report on the Sanitary Conditions of the Labouring Populations of Great Britain* showed the effects of different environments and locations on the health of individual classes. Through statistics and evidence, Chadwick conducted a systematic and unprecedented study of sanitation, examining neighborhoods and streets to identify environmental inequalities. By evaluating class-based death data, he demonstrated that unsanitary environmental circumstances caused more deaths in the working class. Chadwick was adamant that "the attack of fever precedes the destitution, not the destitution the disease."[28] He also calculated that the average age of death was 17 for labourers in Manchester and 52 for gentry in Rutlandshire. Additionally, labourers in Rutlandshire had an average age of death twice that of Manchester labourers. The average age of death for the Manchester gentry was the same as that of the poorest class of Rutlandshire laborers.[29] Chadwick used a statistic called the "average period of life," which was the average age at which a class of people died.[30] By following this approach, Chadwick demonstrated that working class people living in an urban area had the lowest life expectancy, but the main point was that location affected mortality the most. His data showed that members of the same class in the same place died at different ages, depending on their living conditions.[31] Additionally, he found that cottages on hills exposed to

[25] Underwood, "The Centenary of British Public Health: Rise of Health Legislation in England and in London," 890.
[26] Morley, "City Chaos, Contagion, Chadwick, and Social Justice," 68.
[27] Finer, *The Life and Times of Sir Edwin Chadwick*, 160.
[28] Edwin Chadwick, *Report on the Sanitary Condition of the Labouring Population of Great Britain*, ed. M.W. Flinn (Edinburgh: Edinburgh University Press, 1965), 210.
[29] Chadwick, *Report on the Sanitary Condition of the Labouring Population of Great Britain,* 223.
[30] James Hanley, "Edwin Chadwick and the Poverty of Statistics," *Medical History* 46, no. 1 (2002): 21, https://doi.org/10.1017/s0025727300068721.
[31] Hanley, "Edwin Chadwick and the Poverty of Statistics," 27.

uncleanliness harbored as many fevers and diseases as crowded areas.[32] Chadwick's sampling of location and class was unprecedented. It enabled him to conclude boldly, "a high degree of mortality does not invariably belong to the population of all towns, and probably not necessarily to any."[33] Environmental cleanup was critical, and he then turned to engineering to improve sewers, drainage, ventilation, clean water, and overall cleanliness. By relying on statistics and civil engineers, he could focus on prevention rather than cure, setting up the notion of public health and garnering support for solutions that benefit all levels of society, and society as a whole.

Chadwick's report emphasized that filthy environments were to blame, not the poor. Since the environment affected everyone, everyone was responsible, putting forth the idea of the interconnectedness of society's health. Chadwick noted, "It is not these unfortunate creatures only who choose this centre of disease for their living-place who are affected; but the whole town is more or less deteriorated by its vicinity to this pestilential mass."[34] As a result, the annual loss of life due to filth and poor ventilation was greater than the loss from modern wars–a notable finding. The solution to the problem was an overall improvement in cleanliness, as morality and health could not coexist with filth among any community class, since it would harm the whole.[35] Furthermore, when filth was removed, disease almost disappeared.[36] Chadwick's evidence proved this theory. Equal access to sanitation was not just morally right but also necessary for overall public health.

However, a successful public health initiative required the government to intervene and implement infrastructure and laws. As Chadwick noted in a letter to a friend, "I am a zealous advocate for all social improvements and I am therefore an ally of any people by whom improvements would be made. I am therefore a supporter of the existing government."[37] Chadwick wanted to align with any government that would make sanitary improvements. With the help of civil engineers, local public works could operate more effectively and more efficiently.[38] Chadwick asserted that the status quo would continue unless the government seized "sanitary powers" through the legislature.[39] Government authority was crucial for effective public health management, and Chadwick set the stage for the government to assume responsibility and take action.

Public Health Act - A First Step and Catalyst for Progress

Chadwick's work inspired the Public Health Act of 1848, which gave the government centralized authority for clean water, sewer construction, and waste management. The objective

[32] Chadwick, *Report on the Sanitary Condition of the Labouring Population of Great Britain,* 234.
[33] Chadwick, *Report on the Sanitary Condition of the Labouring Population of Great Britain,* 234.
[34] Chadwick, *Report on the Sanitary Condition of The Labouring Population of Great Britain,* 80.
[35] Chadwick, *Report on the Sanitary Condition of The Labouring Population of Great Britain,* 425.
[36] Chadwick, *Report on the Sanitary Condition of The Labouring Population of Great Britain,* 422.
[37] Finer, *The Life and Times of Sir Edwin Chadwick*, 33.
[38] Chadwick, *Report on the Sanitary Condition of The Labouring Population of Great Britain,* 422.
[39] Chadwick, *Report on the Sanitary Condition of The Labouring Population of Great Britain,* 80.

was to use state power to address the problem on a national scale and promote social justice at both national and local levels. The legislation established local health boards as well. Fellow sanitation reformer John Simon noted, "It was an awakening to the duty of caring for the Public Health…the first of our modern legislative endeavours were made to bring Health under the protection of the Law."[40] The government was responsible for bringing medical professionals, statisticians, government officials, and engineers together to solve sanitation problems. While the government now had the authority to guarantee a standard of living, it lacked the enforcement power. Still, the Act established the General Board of Health, which was responsible for advising on public health matters and managing local boards of health, as well as the Health of Towns Commission, which tracked health in British towns. The Health of Towns Commission focused on urban public health and water and sewage issues in the fifty largest metropolitan towns in England and Wales. Each town was examined by a commissioner who worked with a committee of local residents on sanitation infrastructure and public health.[41] PH Holland was active in the Health of Towns Association and served as secretary of its Manchester branch. In a letter to the General Board of Health on December 15th, 1853 Holland wrote "I have ... been taken to see the worst parts of the worst towns in England ... but never did I see anything which could compare with Merthyr ... one of the most strongly marked cases of the evil so frequently observed, of allowing a village to grow into a town, without providing the means of civic organisation. It is the story of laissez-faire carried out to its legitimate conclusion." Merthyr is a town in Wales, and Holland wanted sanitary engineering implemented by the local government–clean drinking water and sewage disposal "before the cholera returns."[42] Holland was keen to draw attention to the perils of the status quo and the pervasive lack of regulations. Beyond commissioners and local committees, citizens also took action to promote the sanitary movement. The Society for the Improvement of the Condition of the Labouring Classes was formed, focusing on solving problems related to filth and disease.

While sanitation reform was in its early stages, but still gaining momentum, Chadwick was under attack for being brash, making mistakes, and acting out of self-interest. Class identities were transforming, and Chadwick's representations of the poor helped him rise up as a member of the middle class. Instead of being named Commissioner of the Poor Law Commission, Chadwick was initially relegated to the position of Secretary of the Commission, in large part due to his status as a middle-class civil servant. He was part of a growing number of middle-class bureaucrats who "managed" the poor, which put him in a position of increasing power. [43]By gathering information for his sanitary report and producing policy, he created ways for the government to intervene in social issues. Indeed, Chadwick simplified the link between

---

[40] Finer, *The Life and Times of Sir Edwin Chadwick*, 515.
[41] Christopher Hamlin, "Health of Towns Association," *Oxford Dictionary of National Biography*, https://doi.org/10.1093/ref:odnb/95366.
[42] Steven Palmer, "From Public Health to the Health of the Public: Modern Public Health Problems Will Not Be Solved by Anything as Simple as Sewers," *BMJ: British Medical Journal* 317, no. 7158 (1998): 550, https://doi.org/10.1136/bmj.317.7158.550.
[43] Priti Joshi, "Edwin Chadwick's Self-Fashioning: Professionalism, Masculinity, and the Victorian Poor," *Victorian Literature and Culture* 32, no. 2 (2004): 353. http://www.jstor.org/stable/25058673.

sanitation and health. However, there were other issues bearing down on the poor, such as meager housing, long working hours, toxic work conditions, and limited education.[44] Sanitary reform would certainly improve public health, but would crime and moral depravity decline? Chadwick believed so, and his work enabled him to present a solution to the complex and messy problem with which the middle classes grappled–how the working class should be handled. Although Chadwick may have been trying to both "protect and police" the poor, and he was solely focused on sanitation as a panacea for many social ills, several of his sanitation-related recommendations were included in the Public Health Acts.[45] Despite the argument that Chadwick was an opportunist who used the poor as a means to gain political power and influence, his larger purpose was to advocate for economic prosperity, and he viewed public health as a necessity to promote industrialization and a free market economy. Chadwick estimated that the annual capital loss in England and Wales from the loss of labor, sickness, and mortality was at least fourteen or fifteen million.[46] Chadwick elevated the importance of urban hygiene and its relationship to public health, driving progress on this vital social issue. Changing class lines and the corresponding shifting power dynamics undoubtedly affected sanitation and public health policymaking. However, Chadwick focused on political economy, which involved public sanitary improvements to water and sewers–developments that would promote economic growth.

Despite his audacious goals, Chadwick made several scientific and tactical mistakes. In 1849, during a cholera outbreak, Chadwick recommended flushing sewage into the Thames, creating a cesspool that would take decades to clean up. Given that the Thames was the primary source of drinking water for London, the risk of contracting cholera was significantly increased due to these actions. The Thames had become an open sewer. In the summer of 1858, the Great Stink occurred–the smell of raw sewage in the Thames prompted Parliament to close. Additionally, although he was correct that the environment led to disease, Chadwick advanced the incorrect miasma theory. He believed disease among the labouring classes occurred because of an impure atmosphere and noted in his report "The various forms of epidemic, endemic, and other disease caused, or aggravated, or propagated chiefly amongst the labouring classes by atmospheric impurities produced by decomposing animal and vegetable substances, by damp and filth, and close and overcrowded dwellings prevail amongst the population in every part of the kingdom."[47] Eventually, Chadwick's scientific mistakes and heavy-handed approach caught up to him. When Chadwick was fired from the General Board of Health in 1854, *The Times of London* said, "We prefer to take our chance of cholera than be bullied into public health."[48] There was

---

[44] Christopher Hamlin, *Public Health and Social Justice in the Age of Chadwick: Britain, 1800–1854* (Cambridge: Cambridge University Press, 2010), 153.
[45] Joshi, "Edwin Chadwick's Self-Fashioning: Professionalism, Masculinity, and the Victorian Poor," 359.
[46] Edwin Chadwick, "Opening Address of the President of Section F (Economic Science and Statistics) of the British Association for the Advancement of Science, Held at Cambridge, in October, 1862," *Journal of the Statistical Society of London* 25, no. 4 (Dec., 1862): 507, https://doi.org/10.2307/2338601.

[47] Chadwick, *Report on the Sanitary Condition of the Labouring Population of Great Britain,* 422.
[48] "This Day in History," Environmental History, Accessed May 9, 2025, https://environmentalhistory.org/this-day-in-history/08-01-1854/.

still significant resistance to government oversight of public health. However, the evidence pointed to the necessity of intervention, and Chadwick kept advocating in spite of the backlash. He continued to expand his knowledge and use of statistics, evaluating the sanitary status of the army in Crimea and India in the 1860s, as well as the Manchester death rate in the 1880s in comparison to other areas.[49] In a 1860 speech, Chadwick stated that the death rate was reduced by one-third through improvements in water supply and drainage within the houses, the removal of cesspools, and the implementation of water-closet apparatus.[50] He remained convinced that the way to maintain social order was through sanitary science and said "To those specially interested in reformatories and the repression of crime, to those specially interested in the repression of pauperism and mendicity, and the relief of sickness and destitution, as well as those interested in the repression of social disorder, I submit, that the primary preventative measure, for which we ought now to labour in common, is to ensure the rising generations a sound mind, having for its necessary foundation a sound body, which must be built up and fortified, and made apt for prolonged and less painful use, by improved physical training under the direction of sanitary science."[51] Despite Chadwick's scientific approach to sanitation, his errors and strong tactics ultimately led to a loss of positioning, eroding his influence and effectiveness.

Fortunately, the sanitation movement had generated significant momentum, and other sanitation reformers persevered, making progress. In 1854, John Snow, a London physician, discovered that cholera was spread by contaminated water. Five hundred deaths within ten days occurred in Soho, London, in what was called the Broad Street Pump Incident. Snow studied the situation and published a book, *On the Mode of Communication of Cholera*, in which he used mapping to help him conclude that cholera was a waterborne infection.[52] One street pump was the cause of the contaminated water, and thus the disease. Snow's use of geographic mapping is a key tool used by epidemiologists today to investigate disease outbreaks such as COVID-19 and Ebola. Furthermore, Snow's work undermined the miasma theory that diseases were spread by noxious air. His findings were instrumental in the development of the germ theory of disease, which asserted that specific diseases are caused by parasitic microorganisms–germs–that invade hosts from external sources, such as contaminated food and water. The germ theory of disease ultimately led to advancements such as vaccines and antibiotics, and strengthened the importance of hygiene measures. In addition, Florence Nightingale improved sanitation in hospitals during the Crimean War, significantly reducing the death rate. Her measures and the data that she published proved that it was disease, not war wounds, that caused most deaths. Her book *Notes on Nursing* conveyed her sentiment that it would be both noble and transformational to use

[49] David Gladstone, "Servicing the Nation's Health: The Later Life of Edwin Chadwick CB." *RSA Journal* 147, no. 5490 (1999): 135, http://www.jstor.org/stable/41378816.
[50] Edwin Chadwick, "Address of Edwin Chadwick, Esq., C.B., as Vice President of the Public Health Section to the General Meeting of the National Association for the Promotion of Social Science, Glasgow, September 29, 1860," *JSTOR Primary Sources* (1861): 7, https://www.jstor.org/stable/60246840.
[51] Edwin Chadwick, "Address of Edwin Chadwick, Esq., C.B., as Vice President of the Public Health Section to the General Meeting of the National Association for the Promotion of Social Science," 39.
[52] John Snow, *On the Mode of Communication of Cholera* (London: John Churchill, 1855), 45, https://archive.org/details/b33093519_0002/page/348/mode/2up.

hygiene as a handmaiden of civilization.[53] Nightingale's use of statistics and evidence, and her emphasis on sanitation, laid the foundation for the principles of modern healthcare.

The Great Stink also provided evidence against the miasma theory and catalyzed engineering ingenuity and Parliament's action on sanitation reform. Due to the stench and its impact on lawmakers, the Metropolitan Board of Works was instructed to initiate work on a new sanitation system in 1859. The board's chief engineer, Joseph Bazalgette, devised plans for an elaborate system–a 82-mile network of main sewers, running parallel to the river, that would intercept both surface water and waste.[54] Basing his calculations on the most densely populated areas of London, Bazalgette derived a number and doubled it, saying: "We're only going to do this once and there's always the unforeseen."[55] The system was completed in 1865 and as a result of Bazalgette's foresight, serves as the foundation of London's present-day sewage system. Society's view of sanitation was transforming into a visionary one–momentum was shifting towards measurable progress and a more orderly society.

Eventually, the sanitation reformers, led by John Simon, also achieved legislative success. The passage of the Public Health Act of 1866 made public health law universal and compulsory, compelling local authorities to improve sanitation. Additionally, provisions were made for inspections and regulations, sewage and drainage, and penalties if rules were not followed. As Simon said "The Act expressly declares it 'the duty' of the local authorities to provide for the proper inspection of their districts, and to proceed for the suppression of whatsoever nuisances should be found existing in them; and, as against the inaction of local authorities, it enacts that, when sewers are not duly provided, or water not duly supplied, or nuisances not duly removed, then, on complaint made, a Secretary of State and the Court of Queen's Bench shall be able to enforce the neglected duty."[56] The Act, also known as the Sanitary Act 1866, formalized the manner in which local governments should handle sanitation matters. The 1875 Public Health Act further mandated that new houses have running water and internal drainage. In 1897, Edward Seaton, Great Britain's Chief Medical Officer, said unsanitary conditions were almost near obsolete in London and in most of the large towns in Great Britain.[57] Progress had not come easily since the British, more than any other European nation, always had "some sort of conscience against filth."[58] However, the beginning of sanitary reform stemmed from the work initiated by Chadwick and the Poor Law Commissioners in 1838. In comparing the periods, Seaton noted that in 1841, the 100,000 people born were reduced to 50,000 at 46

---

[53] Sanna-Leena Rautanen, Antero Luonsi, Henry Nygård, Heikki S. Vuorinen, and Riikka P. Rajala, "Sanitation, Water and Health," *Environment and History* 16, no. 2 (2010): 179, http://www.jstor.org/stable/20723775.
[54] Emily Mann, "Story of Cities #14: London's Great Stink Heralds a Wonder of the Industrial World," *The Guardian*, April 4, 2016, https://www.theguardian.com/cities/2016/apr/04/story-cities-14-london-great-stink-river-thames-joseph-bazalgette-sewage-system.
[55] "Late Great Engineers: Joseph Bazalgette — London's Engineer." *The Engineer*, July 24, 2020. https://www.theengineer.co.uk/content/in-depth/late-great-engineers-joseph-bazalgette-london-s-engineer/.
[56] Edward Seaton, "Evolution of Sanitary Administration in the Victorian Era," *BMJ* 1, no. 1903 (1897): 1636, accessed April 18, 2025, https://doi.org/10.1136/bmj.1.1903.1638.
[57] Seaton, "Evolution of Sanitary Administration in the Victorian Era," 1636.
[58] Seaton, "Evolution of Sanitary Administration in the Victorian Era," 1636.

years old in England and 40 years old in London, while in 1881-90, the number born was only reduced to one-half at the age of 55 and 50 years, respectively.[59] The sanitary reformers in Great Britain had achieved measurable progress in life expectancy and served as an example for other countries to follow.

Legacy of the Sanitary Reformers

Chadwick and the sanitation reformers left a legacy of hygiene in Great Britain that brought London and the rest of England out of a perilous state. Poet Percy Shelley alluded to the pervasive and dire circumstances when he wrote, "Hell is a city much like London." The reformers' experimental, scientific, and systematic approach to sanitation informed how future industrial societies viewed and designed public health. By taking action, the state improved the environment, health, and economic measures for society. All citizens would have access to a sanitary environment regardless of their social status or where they lived because health was viewed as a human right. Disease was not an inevitable reality; it could be controlled through proper sanitation and infrastructure. Espousing these ideals, the sanitation reformers paved the way for the well-being of cities not just in Britain but worldwide. They inspired American politician Lemuel Shattuck's 1850 Report of the Sanitary Commission of Massachusetts, which became the basis of American public health and reinforced Chadwick's emphasis on sanitation and the role of government in public health. Shattuck wrote: "The valuable labors of Edwin Chadwick, Esq., in England, in investigating the sanitary condition of the laboring classes, and the noble efforts which he and his associates have made to induce the government to introduce reforms, deserve the highest praise. Their reports are among the most valuable contributions ever made to sanitary science."[60] By 1869, Massachusetts had established the first State Board of Health, and within nine years, 16 other states followed suit. By the end of the nineteenth century, 40 states and several local areas had established their own health departments. Additionally, fields spanning urban planning to epidemiology have drawn from the work of the sanitary reformers. Life expectancy rose dramatically in areas where public health initiatives took hold.

However, a significant gap still exists among classes, reflecting the need for greater social and environmental justice. Today, in England, life expectancy for men living in the most deprived areas is 73 years, compared with 83 years in the least deprived areas–a difference of 10 years. For women, the difference is 8 years.[61] Disparities can be seen across countries as well. Globally, according to the World Health Organization (WHO), there is a difference of 18 years of life expectancy between high-and low-income countries.[62] Specifically, improper sanitation still

---

[59] Seaton, "Evolution of Sanitary Administration in the Victorian Era," 1640.
[60] Lemuel Shattuck, *Report of the Sanitary Commission of Massachusetts* (Boston: Dutton and Wentworth, State Printers, 1850), 29.
[61] The Health Foundation, "Inequalities in Life Expectancy and Healthy Life Expectancy," *Evidence Hub: Health Inequalities*, February 15, 2025, https://www.health.org.uk/evidence-hub/health-inequalities/inequalities-in-life-expectancy-and-healthy-life-expectancy.
[62] World Health Organization, "Social Determinants of Health," accessed May 8, 2025, https://www.who.int/health-topics/social-determinants-of-health#tab=tab_3.

poses a significant threat in the developing world. According to the WHO, over 1.5 billion people still lack access to sanitation services, including private toilets or latrines. Additionally, 1.4 million people die each year as a result of inadequate drinking water, sanitation, and hygiene.[63] Progress has not benefited the world's citizens equally, and as a result, social and environmental justice remain elusive. Innovation and advancements that improve select lives but not all, undermine real progress.

Perhaps the greatest legacy of the sanitation reformers is their relentless pursuit of social justice alongside progress. Their work catalyzed support and action that cut across class lines, inciting citizens, urban planners, medical experts, and lawmakers alike to collaborate on a common goal and spur meaningful changes in health and well-being. The reformers went beyond laying the foundation for hygiene in industrial Great Britain-they helped redefine the relationship between health, class, and environment in Europe and the world.

[63] World Health Organization, "Sanitation," last updated on March 22, 2024, accessed May 8, 2025, https://www.who.int/news-room/fact-sheets/detail/sanitation#:~:text=Overview,%2Dwater%2C%20sanitation%20and%20hygiene.

Works Cited

Bingham, Paul, Neville Q. Verlander, and M. J. Cheal. "John Snow, William Farr and the 1849 Outbreak of Cholera that Affected London: a Reworking of the Data Highlights the Importance of the Water Supply." *Public Health* 118, no. 6 (2004): 387-394. https://doi.org/10.1016/j.puhe.2004.05.007.

Chadwick, Edwin. "Address of Edwin Chadwick, Esq., C.B., as Vice President of the Public Health Section to the General Meeting of the National Association for the Promotion of Social Science, Glasgow, September 29, 1860." *JSTOR Primary Sources* (1861). https://www.jstor.org/stable/60246840.

Chadwick, Edwin. "Opening Address of the President of Section F (Economic Science and Statistics) of the British Association for the Advancement of Science, Held at Cambridge, in October, 1862." *Journal of the Statistical Society of London* 25, no. 4 (Dec., 1862): 502–524. https://doi.org/10.2307/2338601.

Chadwick, Edwin. "Opening Address of the President of the Department of Economy and Trade, at the Meeting of the National Association for the Promotion of Social Science, Held at York, in September, 1864." *Journal of the Statistical Society of London* 28, no. 1 (1865): 1–33. https://doi.org/10.2307/2338394.

Chadwick, Edwin. "Poor Law Administration, Its Chief Principles and Their Results in England and Ireland as Compared with Scotland." *Journal of the Statistical Society of London* 27, no. 4 (1864): 492–504. https://doi.org/10.2307/2338611.

Chadwick, Edwin. *Report on the Sanitary Condition of the Labouring Population of Great Britain*. Edited by M.W. Flinn. Edinburgh: Edinburgh University Press, 1965.

"This Day in History." Environmental History. Accessed May 9, 2025. https://environmentalhistory.org/this-day-in-history/08-01-1854/

Dickens, Charles. *Oliver Twist.* Global Publishers, 2023. Kindle.

Farr, William. "Opening Address of the President of Section F (Economic Science and Statistics), of the British Association for the Advancement of Science, at the Thirty-Fourth Meeting, at Bath, in September, 1864." *Journal of the Statistical Society of London* 27, no. 4 (1864): 459–78. https://doi.org/10.2307/2338609.

Finer, Samuel Edward. *The Life and Times of Sir Edwin Chadwick*. London: Methuen & Co. LTD., 1952.

Gladstone, David. "Servicing the Nation's Health: The Later Life of Edwin Chadwick CB." *RSA Journal* 147, no. 5490 (1999): 134–35. http://www.jstor.org/stable/41378816.

Hamlin, Christopher. "Could You Starve to Death in England in 1839? The Chadwick-Farr Controversy and the Loss of the "Social" in Public Health." *American Journal of Public Health* 30, vol. 85, 6 (1995): 856-66. doi:10.2105/ajph.85.6.856.

Hamlin, Christopher. "Health of Towns Association." *Oxford Dictionary of National Biography.* https://doi.org/10.1093/ref:odnb/95366.

Hamlin, Christopher. *Public Health and Social Justice in the Age of Chadwick: Britain, 1800–1854*. Cambridge: Cambridge University Press, 2010.

Hanley, James. “Edwin Chadwick and the Poverty of Statistics.” *Medical History* 46, no. 1 (2002): 21–40. https://doi.org/10.1017/s0025727300068721.

The Health Foundation. “Inequalities in Life Expectancy and Healthy Life Expectancy.” *Evidence Hub: Health Inequalities*. February 15, 2025. https://www.health.org.uk/evidence-hub/health-inequalities/inequalities-in-life-expectancy-and-healthy-life-expectancy.

Joshi, Priti. "Edwin Chadwick’s Self-Fashioning: Professionalism, Masculinity, and the Victorian Poor." *Victorian Literature and Culture* 32, no. 2 (2004): 353-70. http://www.jstor.org/stable/25058673.

Late Great Engineers: Joseph Bazalgette — London’s Engineer." *The Engineer*. July 24, 2020. https://www.theengineer.co.uk/content/in-depth/late-great-engineers-joseph-bazalgette-london-s-engineer/.

Henry Mayhew. *London Labour and the London Poor Volume 3*. London: Griffin, Bohn & Company, 1861.

Mann, Emily. “Story of Cities #14: London’s Great Stink Heralds a Wonder of the Industrial World.” *The Guardian*. April 4, 2016. https://www.theguardian.com/cities/2016/apr/04/story-cities-14-london-great-stink-river-thames-joseph-bazalgette-sewage-system.

Morley, Ian. "City Chaos, Contagion, Chadwick, and Social Justice." *The Yale Journal of Biology and Medicine* 80, no. 2 (2007): 61-72. https://pmc.ncbi.nlm.nih.gov/articles/PMC2140185/.

Palmer, Steven. "From Public Health to the Health of the Public: Modern Public Health Problems Will Not Be Solved by Anything as Simple as Sewers." *BMJ: British Medical Journal* 317, no. 7158 (1998): 550–551. https://doi.org/10.1136/bmj.317.7158.550.

Pampel, Fred C. *The Struggle for Public Health*. Baltimore: Johns Hopkins University Press, 2024.

Purdy, Frederick. “The Statistics of the English Poor Rate before and since the Passing of the Poor Law Amendment Act.” *Journal of the Statistical Society of London* 23, no. 3 (1860): 286–329. https://doi.org/10.2307/2338499.

Rautanen, Sanna-Leena, Antero Luonsi, Henry Nygård, Heikki S. Vuorinen, and Riikka P. Rajala. “Sanitation, Water and Health.” *Environment and History* 16, no. 2 (2010): 173–94. http://www.jstor.org/stable/20723775.

Seaton, Edward. "Evolution of Sanitary Administration in the Victorian Era." *BMJ* 1, no. 1903 (1897): 1636-40. Accessed April 18, 2025. https://doi.org/10.1136/bmj.1.1903.1636.

Shattuck, Lemuel. *Report of the Sanitary Commission of Massachusetts*. Boston: Dutton and Wentworth, State Printers, 1850.

Smith, Thomas Southwood. *A Treatise on Fever*. London: Longman, Rees, Orme, Brown, and Green, 1830. https://archive.org/details/b33093519_0002/page/348/mode/2up.

Snow, John. *On the Mode of Communication of Cholera*. London: John Churchill, 1855. https://archive.org/details/b33093519_0002/page/348/mode/2up.

Underwood, Ashworth. "The Centenary of British Public Health: Rise of Health Legislation in England and in London." *The British Medical Journal* 1, no. 4557 (1949): 890-892. Accessed May 1, 2025. https://doi.org/10.1136/bmj.1.4557.890.

Wohl, Anthony S. *Endangered Lives: Public Health in Victorian Britain*. Cambridge: Harvard University Press, 1983.

World Health Organization. “Sanitation.” Last modified March 22, 2024. Accessed May 8, 2025. https://www.who.int/news-room/fact-sheets/detail/sanitation#:~:text=Overview,%2Dwater%2C%20sanitation%20and%20hygiene.

World Health Organization. “Social Determinants of Health.” Accessed May 8, 2025. https://www.who.int/health-topics/social-determinants-of-health#tab=tab_3.

Sacred Rhythms: An Anthropological Study of Drumming, Healing, and Cultural Transmission By Preston Lee

Abstract

The drum is more than just an instrument. Drums have been both a cultural tradition and a therapeutic practice for many societies across many generations. However, the physical composition of drums is often overlooked. This paper examines how the construction methods and the materials of drums influence their therapeutic sound qualities, and how this knowledge has been preserved or revitalized by different cultures over generations. In this paper, I analyzed modern and historical drum-making by different cultures to see how different materials—hardwoods and animal skins—affect the sound properties of drums. Past research has shown that these properties can help support physical, emotional, and spiritual healing. I also examined how drumming relates to cultural identity and how it has been used in therapy. Furthermore, this paper looks at how drumming continuously changes through time, which ensures its relevance in many modern therapy programs and other applications. By conducting case studies on the drums of different cultures, this paper involves discussions on physical and mental health, ethnomusicology, and cultural sustainability. Through these case studies, I also explored different ceremonies, rituals, and traditions connected to drumming. As interest in cultural therapy grows, analyzing the role of drumming in therapeutic practices provides important data for researchers around the world.

Keywords: Drumming, Therapy, Cultural sustainability, Healing, Revitalization, Preservation

I. Introduction

For generations, drums have been used in stories, spiritual practices, and therapy. Across many cultures, the drum has been important in events such as ceremonies, rituals, and even community gatherings. Due to its importance in different societies, the relationship between the construction of drums, their sound properties, and their role in healing continues to be a topic of interest for many. Analyzing this relationship between drum material and culture is crucial to preserving cultural traditions and understanding the drum's usefulness and impact in modern therapeutic practices. Traditional drumming is especially relevant in the world today, where cultural identities are under threat and mental health crises are rising.

What, specifically, is cultural drumming? Cultural drumming refers to the use of drums as a cultural or traditional expression component within a specific community or society (Houston). It measures the rhythms, techniques, and styles of drumming that are deeply connected to the cultural identity and history of a group of people. These drum practices vary widely from region to region and are often passed down from generation to generation. For instance, the Djembe (Western African Drum) originated from the Mandinka people of Western Africa around 1300 AD, and its cultural implications have been preserved ever since (Price). The Djembe was played at group gatherings and rituals, which allowed the members of the Maninke tribe to have spiritual revivals and heal collectively. Hence, cultural drumming can be used in

musical therapy, but it is also a spiritual connector for many groups.

Cultural healing is the process of using traditions and communal practices to foster emotional and spiritual well-being. It is the purpose of many cultural drumming traditions. Past studies have shown that this style of drumming can help reduce stress, improve mental health, and create a sense of belonging for many, particularly in marginalized communities. Examples of this can be seen in research based on indigenous cultural drumming. Specifically, one study shows how indigenous drumming rhythms in Africa serve to reduce anxiety and stress (Núñez). The rhythms are used to channel energy, which researchers believe can lead to improved overall health. In African culture, drum rhythms are extremely important because they act as medicinal instruments and ceremonial focal points across Indigenous communities (Núñez). Another study, based in the Western Hemisphere, reinforces Núñez's ideas by finding that these therapeutic outcomes, often called "spiritual healing," are associated with cultural drumming (Houston). In this study, Houston finds the newer implications of cultural drumming and its effects on modern culture. Houston reveals that while there are many physiological factors, such as exhaustive dancing and auditory driving, that may play a role in the drum's overall therapeutic effect, symbolic processes (cultural drumming) provide the necessary framework to facilitate healing.

Current research in music therapy has shown how drumming circles create safe spaces for expression and community bonding. For example, a drum ritual that originates from Furukawa, Japan, functions as an important way for communities to bond together. They bond through shared conflict resolution and collective participation (Schnell). Taking place during the annual Shinto shrine festival, this event features a large barrel-shaped drum that is carried through the streets on a grid platform. The community leaders ride on the platform while groups of young adults from different neighborhoods rush out to physically attack it as it passes by. Historically, this chaotic scene was a way for Japanese citizens to express unresolved tensions and even settle disputes within the town. However, over time, especially during Japan's modernization in the late 1800s, the ritual evolved into a symbolic challenge against institutional authority, with occasional politically motivated violence (Schnell). Although this ritual demonstrates violence and conflict, the culture is enjoyed by many people and constructs a close community within Furukawa. These cultural drumming traditions allow groups to have a comfortable space for bonding between community members and their cultural origins.

This paper will examine the materials and construction techniques of drums across four different cultures and analyze how the physical attributes shape each of their therapeutic effects. This paper will also analyze how drumming traditions and methods of construction have been preserved or revitalized over generations. Through the analysis, it becomes clear that drumming holds many implications that go far beyond being just an artifact. Drumming is a living tradition that continues to heal communities today. By analyzing the physical components of these drums, along with the way the materials and construction influence the sound, we will be able to better understand why drumming functions to heal and makes such a significant impact on many groups. Through this, we may attain a deeper understanding of the ways drumming is used, how it works, and how it can be used in the future.

II. The Intersection of Drum Construction, Cultural Preservation, and Healing

In order to understand how cultural healing through drumming works, it is important to first analyze the construction of the drum, how the construction influences its therapeutic sounds, and the ways these methods have been protected over time. This helps us address two key aspects: the physical and acoustic properties of drums, and their cultural implications in events such as ceremonies.

There are three levels of analysis that can be used to determine the cultural implications of indigenous drumming:

The first is on a therapeutic level. This level helps identify the specific materials, such as types of wood or animal skins, and construction techniques, such as tensioning methods or drum shapes, that are used. These concepts can help explain how different drums produce distinct vibrations and frequencies that can impact mental and physical health. For instance, there are certain frequencies that might align with the human body's natural rhythms, which may be helpful for relaxation, stress relief, or even physiological healing. An example of this can be seen in a study based on drumming for people with cancer. In the experiment, Harmon finds that drumming can effectively help reduce stress and promote feelings of happiness, making it an important aid in coping with illnesses such as cancer (Harmon). Harmon continues that cancer patients have often struggled with the ability to express emotions due to their diagnosis or the demands of needed treatment. However, therapeutic drumming can be a form of creative expression for many patients and can help them reclaim their sense of agency and autonomy (Harmon). Drumming as an expression can give a really impactful boost to the well-being of individuals during an especially challenging period of their lives.

Secondly, on a cultural level, studying how these construction methods are preserved allows us to understand the role drums play in maintaining group identities. This level also gives us insight on how to preserve or transfer that cultural knowledge across generations. This includes how specific techniques and materials are chosen not just for sound quality, but also for their symbolic meaning within a culture. For example, the Wadaiko drum, which is traditionally crafted from wood, has a deep cultural history in Japan, as it was very commonly used in rituals and festivals to unite communities and invoke spiritual healing (Ono et al.). Today, the modern application of combining that wood with plastic allows for better sound clarity while preserving the drum's original tones. This is just one of many examples that reveal how studying the construction methods of cultural drums will continue to play a vital role in cultural revitalization.

Lastly, on an applied level, understanding these processes allows for opportunities to use therapeutic drumming practices in modern-day society while respecting their origins. As we have seen in the 20th to 21st century, this could involve creating programs that address issues such as mental health. For example, a study on group drumming interventions showed significant improvements in anxiety, depression, and social resilience among mental health service users (Fancourt et al.). The participants of this study engaged in ten weeks of group drumming and experienced decreases in depression and anxiety levels, and experienced increased mental health levels. The research also showed biological benefits, including a shift towards a more

anti-inflammatory immune profile. These studies further support the potential of using drumming for therapy.

III. Background

Drumming is one of the oldest forms of cultural expression, with evidence of it dating back thousands of years. From the hollow log drums of early hunter-gatherers to the newly designed instruments of modern groups, drumming has evolved and served in different societies. Today, drumming continues to be a topic of interest across many fields—anthropology, ethnomusicology, and music therapy—as well as among researchers who study its cultural and therapeutic impacts.

For the purposes of this paper, these are some key concepts to know:

Cultural revitalization, in the context of cultural drumming, refers to organized efforts by community members to reconstruct traditional drumming practices with the goal of strengthening the community's identity and resilience (Shea et al.). By reconnecting with these cultural practices, modern communities are able to link the past to the present through shared drumming as a therapeutic tool. This process has been linked to initiatives that envision ways in which cultural practices can be reintegrated into contemporary society (Isaac).

Then, preservation ensures that traditional drumming techniques and the knowledge behind them are maintained for future generations (Shea et al.). By preserving these practices, we protect cultural heritage and continue the legacy of using drumming for therapeutic purposes. Preservation allows us to pass down valuable cultural knowledge, which can be applied to society today. An example of preserving traditional drumming practices is the Minneapolis American Indian Center's (MAIC) drum and dance sessions. These gatherings help to reconnect individuals with their ancestors' cultures and ways of life. As drum instructor Tony Frank noted, the sessions assist those who "may feel lost inside" without connections to their heritage, helping them find their way back (Dell'Orto).

Social resilience refers to the ability of individuals or communities to adapt to, recover from, or thrive in the face of challenges (Keck and Sakdapolrak). In the context of drumming, social resilience can be enhanced by establishing unity through communal drumming practices, which contribute to a community's ability to cope with stress and other setbacks; This concept includes coping capacities—the ability of social actors to overcome adversities—and adaptive capacities—the ability to learn from past experiences and adjust to future challenges (Keck and Sakdapolrak). A study by Ho et al. (2011) reveals evidence supporting the role of communal drumming in enhancing social resilience among low-income, primarily Latino children. Participants engaged in group drumming sessions led by school counselors, which resulted in significant improvements in multiple domains of social-emotional behavior, including reductions in depression and attention problems.

IV. Current State of Research

Both the therapeutic and cultural importance of drumming have been well-studied and researched. Many academics have explored how unique cultural practices like drumming create

community recovery. A study that was conducted between January and August of 2012 focused on the Community Drumming Networks' (CDN) social impact, which pointed out how drum circles are a platform for people of diverse ethnic backgrounds to unite (Onishi). This study demonstrated how drumming could create racial harmony between distinct groups. The groups that participated throughout the whole eight-month experiment stated that joining drum circles with other cultures helped them strengthen their community connections and also helped them build up their social resilience. However, these circles not only served as a form of expression for these groups, but they also served as a way for communal healing and unity. Ultimately, this study displays drumming's function of connecting people and agrees with the understanding of its positive effects on community well-being.

Another important example of drumming being used in recovery is a study conducted by Bittman et al. (2001), which explored the physiological and psychological effects of group drumming on people with chronic stress and other related conditions. The researchers designed a controlled experiment where their participants engaged in structured drumming sessions over six weeks. The data gathered indicated that drumming heavily reduced cortisol levels—a biomarker of stress—while also improving the participants' moods. The study's methodology was unique as it used a combination of emotional and biological assessments, such as immune function markers, to measure the impact of drumming on a person's overall health. Because the findings clearly reveal reductions in stress and improvements in emotions, this study further supports the claim that drumming is an impactful resource to address mental and physical conditions.

These studies demonstrate how drumming helps people from different communities cope with stress, rebuild identities, and connect with others. Research on group drumming has repeatedly shown its effectiveness in treating mental health concerns such as anxiety and depression. Similarly, music therapy research has also revealed how the rhythmic and communal nature of drumming promotes relaxation, cognitive function, and social bonds (Crawford et al.). Additionally, the buildup plus synchronization of rhythm in group drumming has been linked to the release of endorphins and the regulation of brainwave activity, which can lead to inner balance (Crawford et al.). Thus, the use of drums goes well beyond the creation of sound for entertainment.

V. Drumming Across Cultures (Analysis)

Drumming traditions vary widely between cultures, which can inform us about the differences in beliefs and practices within societies. By analyzing four case studies—the Baganda, Acholi, Djembe (West Africa), and Native American Frame Drum—we can understand how the materials and construction methods of drums affect their tone and sound. Through this, we may also understand how drumming is preserved and what its therapeutic applications are.

To provide historical background for the following case studies, this section is an overview of some important details:

Drums have and continue to be an important part of the identities of the Baganda and Acholi tribes in Uganda. The Baganda have always used royal drums as a sign of authority and to connect with their ancestors spiritually (Kafumbe). The two most famous royal drums that

they use are the Mujaguzo and Kawuugulu. In contrast, the Acholi have used war drums to unite warriors and send messages, also as described by Kafumbe. Their war drums were made with unique materials that were chosen for their own specific qualities. While the Baganda still use their royal drums for the same purpose, the Acholi's use of drums has evolved over time.

In West Africa, the djembe is a very popular traditional drum that is made from different kinds of wood. The primary purpose of the djembe among tribes has been for either entertainment or communication uses, and it continues to serve that purpose today. Lastly, the Native American Frame Drum is a commonly used instrument by many Native tribes, although its specific purpose may vary from tribe to tribe. However, this study will focus on its use within a very specific Native American ceremony. These drums have been played in such ceremonies for centuries.

A: Baganda Tribes

The Baganda tribe, native to East Africa, uses many different kinds of wood to achieve different sounds. The drums contain different sounds for different occasions. For deep tones with echoes, they use two-ply and three-ply plywood. This gives the drum a very dark tone. When they want sharper and clearer sounds, they find dense hardwoods like oak and mahogany. Denser wood provides a brighter quality that is ideal for ceremonial performances. On the other hand, when they want a warmer tone, softer woods, such as cedar or particular indigenous species, are used to create a more soothing sound. The choice of wood is not only a matter of sound, but it also carries a deep cultural meaning. In Bagandan culture, different trees and specific woods are believed to symbolize spiritual and royal power (Kafumbe). Kafumbe finds that due to the uniqueness of the materials required to tune and maintain the royal drum set, the Baganda people heavily rely on the guidance of ancestral spirits to locate the best resources.

Once the wood has been chosen, Baganda drum makers follow a strict construction process to shape the instrument's sound and durability. The makers must follow a specific order to get the desired sound. According to Kafumbe, the makers use chisels and adzes to carve out the drum body from a block of wood, and they hollow out the center using only their hands and tools (Kafumbe). Additionally, the thickness of the drum walls is carefully controlled because thinner walls allow for more vibration and louder sounds. On the other hand, thicker walls create a more muted tone. After the carving process, the wood is left outside to dry. Drying the wood prevents the drum from cracking later on and helps the drum produce a cleaner sound over time (Kafumbe). Some drum makers also add internal grooves or slight tapers inside the shell to help direct the flow of sound waves (Fleder). The internal grooves or tapers inside the shell may ultimately enhance the drum's performance if added. The drum is then smoothed out and sealed to protect it from moisture. This finishing process allows it to stay in good condition during use in both ceremonies and outdoor events.

As revealed, the entire process of constructing Baganda drums is very well thought out by their craftsmen. However, the process of making the drum is overlooked by many outsiders. But without the specific materials and construction process, Bagandan drums would not be able to produce the sound or serve the purpose they are known for. For example, a famous drumming

pattern that the Baganda people often use is the Gahu rhythm. Although the rhythm has a strong presence in Baganda traditions, it originated in West Africa. Baganda drum makers purposefully shape their instruments to follow this rhythm. This rhythm follows a slow and steady tempo, and it features a call-and-response structure. For example, one drum may lead with a pattern like "dum, dum, da-da-dum", and another will respond with "da-da-da-dum, dum-dum" (Fleder). This interactive style of drumming builds communication among drummers and dancers, which allows for important cultural performances that engage both musicians and audiences (Nketia).

The Gahu rhythm is an important part of Bagandan society, especially during royal ceremonies. It's used to show the authority and power of the monarchy. In the past, drumming helped reinforce social structures, and certain rhythms were specially made just for kings, chiefs, and other important figures (Roscoe). When Gahu is played at royal events, it represents prestige. During the Gahu, there is also a back-and-forth tempo between drummers. This back-and-forth communication makes it feel like the drums are talking to each other. The overall tempo is meant to reflect the relationship between rulers and their people (Nannyonga-Tamusuza).

During these events, both the rhythm's speed and intensity can change depending on what the purpose of the gathering is. For example, at a funeral or ritual for a high-ranking leader, the drumming might be a little slower to show signs of respect and mourning (Kubik). But during celebratory events like coronations, feasts, or victories, the rhythm becomes more upbeat to encourage dancing and celebration among community members (Nketia). These changes in rhythm help show the emotions and importance of each moment in these events (Nannyonga-Tamusuza).

Besides being used in royal events, the Gahu rhythm also shows up in other parts of Baganda culture. It's used in many traditional dances, storytelling, and community celebrations. Because it's used in so many different ways, the Gahu rhythm has become a key part of Baganda life. But it's important to remember that the rhythm depends on the drum itself. Without the right materials and the way the drums are made, the Gahu rhythm wouldn't work the way it's meant to.



Fig 1: Bagandan Royal Drums (Singing Wells)

B: Acholi Tribes

Unlike the Baganda, the Acholi used to make their drums with buffalo skin, since they were originally hunters (Girling). Buffalo hide was a good choice because it was thick and strong, which helped create powerful sounds that worked well for rituals and community gatherings. But over time, the Acholi moved away from using buffalo skin. This shift happened for a few reasons, including changes in their lifestyle, fewer buffalo being available, and changes in how their society was structured (Behrend).

As the Acholi shifted from living as hunter-gatherers to focusing more on farming (agrarian lifestyle), they began to depend less on buffalo. Since buffalo hunting became less common, it was harder for them to find buffalo hides, which made it difficult to keep up their traditional drum-making methods (Girling). Additionally, colonial-era rules about wildlife conservation and hunting placed even more limits on access to buffalo skins, so the Acholi had to change how they made their drums (Behrend). Over time, they started using other materials, especially cow and goat hides, which were easier to attain since they were raising more livestock (Otim).

In addition to changing the types of skins used for drumheads, the Acholi, like the Baganda, also made their drums from tree trunks. The bul drums, which play a big role in Acholi music and dance, were usually carved from hardwoods like mvule (Milicia excelsa) and mahogany. These woods were chosen because they were strong, lasted a long time, and could produce deep and rich sounds (Girling). Drum-making in Acholi society is considered to be a skilled craft. Artisans need to know how different woods affect sound in order to select the perfect one. After choosing the right material, the craftsmen would then start to carve and shape the drums by hand.

The size and height of the bul drums affected both their pitch and their role in Acholi music. Smaller drums had higher-pitched sounds and were mostly played in the Bwola dance, a royal performance traditionally held to honor chiefs and other important leaders within Acholi society (Otim). Larger drums made deeper sounds and were often used to show power, signal changes, or draw attention during ceremonies and community gatherings (Behrend). When the big and small drums were played together in sync, they created a layered and polyrhythmic style that is a key part of Acholi music.

Even though the Acholi changed the materials they used, drummers still followed a careful process to make sure their drums were authentic. The most important sign of quality was always the richness and depth of the sound, so musicians would test each drum before using it in a performance. This tradition has been passed down through generations. The preservation of this tradition helps keep Acholi drumming alive even as changes in the environment and society affect how the drums are made (Otim). Drumming is still a key part of Acholi identity, whether it's for social, spiritual, or entertainment purposes. Today, these traditions have been preserved and adapted for new purposes, such as modern music performances, cultural festivals, and efforts to protect their heritage (Behrend). The Acholi keep their musical legacy alive by holding on to their drumming styles and instruments. This tradition reflects both their long history and their

resilience.

Ultimately, while the Baganda's royal drums are tied to authority and spiritual connection, the Acholi's drumming traditions grew out of societal changes. These changes show their resilience and ability to adapt. Even though their materials and techniques have changed over time, drumming is still an important part of Acholi identity.



Fig 2: Acholi Bul Drums (FOBE Travels)

C: The Djembe

In West Africa, the djembe is one of the most recognizable drums. It has a goblet-shaped body that is carved from a single piece of hardwood, with a tightly stretched goat-skin head on the top. These parts work together to give the drum a wide range of tones, which is why it's used for both music and communication. Traditionally, carvers used hardwoods like Lenke, Djalla, or Iroko, since these woods make the sound more resonant and the drum more durable (Charry). Once a tree is selected, the outer bark is stripped and the trunk is roughly shaped before finer carving begins. The outside is smoothed for a better appearance, while the inside is shaped to control how the sound will travel through the drum. The drumhead is made from goat skin, which can produce three main tones: bass (deep), tone (mid-range), and slap (high and sharp). The thickness of the skin changes the sound. Thinner skins give a sharper sound, while thicker skins make a warmer and fuller tone (Charry).

The djembe is also a symbol of unity for many communities. It plays an important role in ceremonies such as harvest celebrations, spiritual events, and rites of passage—ceremonies that mark important life transitions like birth, marriage, or adulthood (Price). The rhythms can act like a language, sending messages within a gathering or across distances. When people drum together, dancers respond to the beat, and others join in by clapping or chanting. These shared experiences give the djembe a lasting place in West African culture and make it a key part of community life (Price).

The formula for making a djembe has been passed down for generations in many West African tribes, including the Mandinka and Bambara. This has helped keep both the craftsmanship and the playing techniques alive. In the past, this knowledge was shared directly between master drum makers and apprentices, often within the same family. Young learners

would spend years observing, assisting, and practicing under guidance before being trusted to make a drum on their own (Charry). In many tribes, teaching the process to the next generation also meant preserving the traditions, ceremonies, and rhythms tied to it.

Today, efforts to preserve the tradition include drum-making workshops, cultural festivals, and training programs in Africa and other parts of the world (Charry). These activities keep the tradition going, even with the effects of global commercialization. Due to high demand, some modern makers use synthetic drumheads and factory-made frames instead of traditional materials, which can alter the sound quality of the drum. Even so, many West African artisans still follow the old methods so they can keep the djembe's traditional sound and craftsmanship.

One of the most well-known ways the djembe is played is through a call-and-response rhythm (Charry). In this style, one drummer plays a short pattern, and the others respond with a pattern of their own. This back-and-forth creates a conversation between the drums and often gets dancers and the audience involved, making it an important part of many West African performances.

The djembe is also very popular due to its wide range of applications. Specifically, it is very famously used in music therapy. Research shows that playing the djembe in groups can help lower stress and anxiety by syncing brainwave activity and reducing cortisol levels, which makes it useful in mental health programs and trauma recovery (Koen et al.). Its wide range of sounds lets therapists design sessions where the different tones—bass, tone, and slap—can match or even change how a patient is feeling. The call-and-response style that's common in djembe playing gets people more involved and makes them feel connected during therapy. Since some rhythms are simple to pick up, beginners are able to join in. It's also used in neurological rehabilitation to help stroke patients and people with neurodegenerative diseases improve motor coordination and thinking skills (Koen et al.).



Fig 3: West-African Djembe Drum (Brixton House)

D: Native Tribes

Among many Native American cultures, the frame drum is a very important part of their spiritual and ceremonial life. The drum is made with a wooden frame and a rawhide head, and the hide is laced traditionally so that the tension changes how the drum sounds (Bucko). The hide is usually from animals like deer or elk because they are durable and give the drum a unique tone. To make one, the wood is shaped into a round frame, then the hide is soaked to soften it

before being pulled tightly across. As it dries, it naturally tightens and creates the tension needed for the sound. Frame drums come in different sizes, and the size changes the way they sound. Larger drums have deeper tones that can be heard from far away, which works well for outdoor ceremonies. Smaller drums make higher-pitched sounds that stand out more in fast-paced healing rituals (Bucko). In some traditions, the drum is played with a beater, while in others, people use their hands to have more control over the tone and rhythm.

After making the foundation of their drums, many Native American communities decorate their drumheads with sacred symbols and natural pigments to show their connection to their ancestors and the earth (Bucko). These designs might include patterns, animals, or colors that have special spiritual meaning in the tribe. Decorating the drum is sometimes just as important as playing it because it's a way to honor spirits and ask for guidance. The drum itself is seen as a bridge between the physical world and the spiritual one. In ceremonies, the beat can represent the heartbeat of the earth. This sound represents life, balance, and the link between people and nature. For many Native American groups, hearing that rhythm is a powerful reminder of who they are and how they are connected to the land.

One example of connecting the physical and spiritual can be seen in the Lakota people. They include drumming in their sweat lodge ceremonies, which are purification rituals meant to improve emotional, spiritual, and physical health. The lodge is a small dome made from willow branches tied together and covered with hides or blankets. This creates an enclosed space that traps heat during the ceremony. Inside the lodge, the drum's steady beat is played to match what is believed to be the heartbeat of the earth, helping people stay focused and enter a meditative state (Bucko). The small space is heated by stones that have been warmed in a fire, and water is poured over them to create steam. The intense heat causes sweating, which is said to cleanse the body of toxins, while the drum's sound and ceremonial songs guide people through the process. In Lakota tradition, this experience is a way to pray, reflect, and reconnect with the community and the natural world. When people "sweat" together, they go through personal purification and also feel unified with others. This connection, along with the ceremony's importance to their culture, has helped the sweat lodge stay a strong part of Lakota identity even with outside pressures over the years. It shows their resilience and their dedication to keeping these healing practices alive for future generations (Bucko).

People have been using frame drums in healing and spiritual ceremonies for a long time, and many believe that they really can help with health and well-being. Recent studies have supported these beliefs. Researchers studying Native drumming therapy in their own communities found that it can help people recover from trauma, treat PTSD, and even support addiction recovery (Neher). They've also found that frame drumming can improve heart rate variability and lower symptoms of stress-related problems. This fits with what other research says about how music therapy can help the body (Neher).

Additionally, in many Native American communities, people work hard to preserve both the frame drum and the traditions that come with it, like the sweat lodge ceremony, from fading away. Elders pass down skills such as choosing the right materials, shaping the wooden frame,

and lacing the rawhide so the drum keeps the same sound it's had for generations (Bucko). Some of this knowledge is shared at workshops, cultural events, and community gatherings (Bucko). During these events, teachers also talk about which woods and hides to use—some tribes pick certain materials based on the sound, while others choose them for cultural value. Even though newer materials are easier to find, many drum makers still stick to the old ways.

After all these efforts to preserve the sweat lodge, it has still faced periods of decline. However, the ceremony has made multiple comebacks. Younger members actively try to learn the meaning behind it, as well as the songs and rhythms, from elders and cultural leaders. This keeps an important way of life alive and connected to the land. Even with government pressure and past bans on Native ceremonies, many communities have still held onto these traditions and found ways to keep them going (Bucko). Being able to do that shows the strength and resilience that has helped Native cultures last for generations.



Fig 4: Native Sun-Designed Frame Drum (Native Art Shop)

VI. Modern Drumming Practices

In recent years, drumming has gotten more attention for how it can be used to help in different settings. Many studies have found that group drumming can help improve mental health, especially by lowering anxiety and depression. In one particular study, Perkins et al. analyzed how group drumming affected people who use mental health services and their caregivers. They found that the drumming sessions helped people feel more socially connected and made them feel more accomplished, more in control of what they were able to do. The sessions also gave participants a way to express themselves without words and created a supportive space where they could connect with others and grow personally (Perkins et al.).

An example of a program that utilizes drumming for therapy is HealthRHYTHMS. This program was created by experts in both brain science and music therapy, working together to design group drumming sessions that boost emotional and physical health (Remo). They focus especially on helping people who are recovering from trauma or dealing with long-term stress. Many participants say the sessions make them feel more involved in life and better able to manage their health (Perkins et al.).

Drumming has also been used as an extra form of therapy in medical settings. In one

study, Ho et al. looked at how group drumming affected social and emotional behavior in kids from low-income families. They found that after taking part, the children showed fewer signs of withdrawn behavior and were better able to focus in school. This suggests that drumming can help children gain confidence and strengthen their relationships with others, leading to overall social growth (Ho et al.).

Aside from mental health benefits, drumming can also have physical benefits. Past research has shown that rhythms can lower blood pressure and help improve motor skills. It can also help with chronic pain because it lets the body release endorphins, which are natural chemicals that help block pain (Bitmann et al.). Because of these effects, drumming has become a way for people to feel healthier and manage stress in their daily lives.

A lot of these more recent applications of drumming have often gotten their inspiration from traditional practices. This shows how cultural knowledge can be kept alive while also being adapted for today. By bringing drumming into modern therapy programs, people can respect its history while also using it to solve current health problems. Whether it's in hospitals, rehab programs, or community events, drumming still plays an important role in helping people heal and connect with each other (Perkins et al.).

VII. Conclusions/Future Directions

Drums have been a vital part of human life for centuries, carrying stories, traditions, and healing practices that are passed down from one generation to the next. The materials that are used to make the drums are purposefully chosen for their sound and for how valuable they are to different cultures. Around the world, drumming has been used for settings that include emotional release, building friendships, and even helping people recover physically. From West African celebrations to Native American ceremonies to current therapy programs, drumming helps communities stay connected and preserves their heritage. In the future, it will be important for modern uses of drumming to respect and protect this cultural knowledge so that it can be preserved for upcoming generations.

Ultimately, this paper highlights how drumming functions not only as a preserved or revitalized cultural tradition, but also as an effective way for supporting health. Knowing the drum's value across different cultures and the effects it has on the human body can help bring it into healthcare, schools, and different programs. Currently, drumming has already been used to help people handle stress, anxiety, and other emotional challenges. In the future, researchers could study how different cultures share and adapt drumming to help people heal. It would also be meaningful to look into how drumming changes brain activity, since that could help people better understand its role in healthcare. Future research in anthropology could go further by diving deeper into how drumming shapes cultural identity, keeps communities strong, and preserves traditions for the future.

Acknowledgements

I express my deepest gratitude toward Professor Kasey Diserens Morgan for her invaluable support throughout the entire research and writing process.

Works Cited

Bittman, B. B., Berk, L. S., Shannon, M., Sharaf, M., Westengard, J., Guegler, K. J., & Ruff, D. W. (2001). Composite effects of group drumming music therapy on modulation of neuroendocrine-immune parameters in normal subjects. Alternative Therapies in Health and Medicine, 7(1), 38-47. https://pubmed.ncbi.nlm.nih.gov/32277749/

Brixton House. (n.d.). African djembe drumming. https://brixtonhouse.co.uk/shows/african-djembe-drumming-2/

Bucko, R. A. (1998). The Lakota ritual of the sweat lodge. University of Nebraska Press. https://doi.org/10.1525/esr.1998.21.1.108

Charry, E. (2000). Mande music: Traditional and modern music of the Maninka and Mandinka of Western Africa. University of Chicago Press. https://doi.org/10.1353/ral.2001.0061

Crawford, P., Lewis, L., Brown, B., & Manning, N. (2013). Creative practice as mutual recovery in mental health. Mental Health Review Journal, 18, 55–64. https://doi.org/10.1108/MHRJ-11-2012-0031.

Dell'Orto, G. (2024). Native Americans in Upper Midwest protect their drumming tradition. Associated Press News. Retrieved from https://www.ap.org/news-highlights/spotlights/2024/safeguarding-the-heartbeat-native-americans-in-upper-midwest-protect-their-drumming-tradition/.

Fancourt, D., Perkins, R., Ascenso, S., Carvalho, L. A., & Steptoe, A. (2016). Effects of group drumming interventions on anxiety, depression, social resilience, and inflammatory immune response among mental health service users. PLOS ONE, 11(3), e0151136. https://doi.org/10.1371/journal.pone.0151136.

Fleder, M. (n.d.). Music of Africa final paper. https://ocw.mit.edu/courses/21m-293-music-of-africa-fall-2005/abbbcec33a27eade4212b46efcd49078_gahu.pdf

FOBE Travels. (2019, August 18). A tale of generations: The Acholi traditional handicrafts rooted in culture. https://fobetravels.wordpress.com/2019/08/18/a-tale-of-generations-the-acholi-traditional-handicrafts-rooted-in-culture/

Girling, F. K. (1960). The Acholi of Uganda. Her Majesty's Stationery Office. https://archive.org/details/acholiofuganda0000fkgi

Harmon, J., & Arpajian, A. (2020). Restorative rhythms: Drumming as healing. World Leisure Journal, 62(1), 67-80. https://www.tandfonline.com/doi/full/10.1080/16078055.2019.1611629

Ho, P., Tsao, J. C. I., Bloch-Salisbury, E., & Zeltzer, L. K. (2011). The impact of group drumming on social-emotional behavior in low-income children. Evidence-Based Complementary and Alternative Medicine, 2011, Article ID 250708. https://doi.org/10.1093/ecam/neq072

Hutson, S. R. (2000). The rave: Spiritual healing in modern Western subcultures. Anthropological Quarterly, 35-49. https://voidnetwork.gr/wp-content/uploads/2016/09/Rave-as-spiritual-healing-in-Modern-Western-Subcultures-by-Scott-R.-Hutson.pdf

Isaac, G. (2022). Being present and bearing witness: Talking about cultural revitalization. International Journal of Heritage Studies, 27(6), 575–588. https://www.researchgate.net/publication/362462156_Being_present_and_bearing_witness_talking_about_cultural_revitalization_programming_in_museums

Kafumbe, D. (2011). The Kawuugulu royal drums: Musical regalia, history, and social organization among the Baganda people of Uganda (Doctoral dissertation, The Florida State University). Florida State University Libraries. https://repository.lib.fsu.edu/islandora/object/fsu:254183

Keck, M., & Sakdapolrak, P. (2013). What is social resilience? Lessons learned and ways forward. Erdkunde, 67(1), 5–19. https://doi.org/10.3112/erdkunde.2013.01.02

Koen, B. D., Lloyd, D., Barz, G., & Brummel-Smith, K. (2008). The Oxford handbook of medical ethnomusicology. Oxford University Press. https://api.pageplace.de/preview/DT0400.9780199714155_A23608279/preview-9780199714155_A23608279.pdf

Kubik, G. (2010). Theory of African music, Volume II. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/T/bo8648201.html

Nannyonga-Tamusuza, S. (2005). Baakisimba: Gender in the music and dance of the Baganda people of Uganda. Routledge. https://cmc.marmot.org/EbscoAcademicCMC/ocn870272770

Native Art Shop. (n.d.). 16" drum sun design. https://www.nativeartshop.ca/product/16-drum-sun-design/

Neher, A. (1961). Auditory driving observed with scalp electrodes in normal subjects. Electroencephalography and Clinical Neurophysiology, 13(3), 449–451. https://doi.org/10.1016/0013-4694(61)90014-1

Nketia, J. H. (1995). The music of Africa. W. W. Norton & Company. https://www.cambridge.org/core/journals/yearbook-of-the-international-folk-music-council/article/abs/j-h-kwabena-nketia-the-music-of-africa-new-york-w-w-norton-company-1974-x-278-pp-maps-photos-and-musical-examples-appendices-bibliography-and-index/6BB3E26D810F2EB1B9EE2733992BB109

Núñez, S. (2016). Medicinal drumming: An ancient and modern-day healing approach. NeuroQuantology, 14(2). https://www.researchgate.net/publication/303635377_Medicinal_Drumming_An_Ancient_and_Modern_Day_Healing_Approach

Onishi, P. C. (2014). Drumming for community building: The development of the Community Drumming Network (CDN) and its impact in Singapore society. International Journal of

Community Music, 7(3), 299-317. https://www.researchgate.net/publication/272102772_Drumming_for_community_building_The_development_of_the_Community_Drumming_Network_CDN_and_its_impact_in_Singapore_society

Ono, T., Takahashi, I., Takasu, Y., Miura, Y., & Watanabe, U. (2009). Acoustic characteristics of Wadaiko (traditional Japanese drum) with wood-plastic shell. Acoustical Science and Technology, 30(6), 410-416. https://www.jstage.jst.go.jp/article/ast/30/6/30_6_410/_pdf

Perkins, R., Ascenso, S., Atkins, L., Fancourt, D., & Steptoe, A. (2018). Promoting well-being through group drumming with mental health service users and their carers. International Journal of Qualitative Studies on Health and Well-being, 11, 31315. https://pubmed.ncbi.nlm.nih.gov/29989487/

Price, T. Y. (2013). Rhythms of culture: Djembe and African memory in African-American cultural traditions. Black Music Research Journal, 33(2), 227-247. https://www.academia.edu/8544757/Rhtythms_of_Culture_djembe_and_African_memory_in_American_American_cultural_traditions

Roscoe, J. (1911). The Baganda: An account of their native customs and beliefs. Macmillan. https://doi.org/10.1017/CBO9781139051071

Shea, H., Mosley-Howard, G. S., Baldwin, D., Ironstrack, G., Rousmaniere, K., & Schroer, J. E. (2019). Cultural revitalization as a restorative process to combat racial and cultural trauma and promote living well. Cultural Diversity and Ethnic Minority Psychology, 25(4), 553–565. https://doi.org/10.1037/cdp0000250.

Singing Wells. (n.d.). The Entenga drums (part 1). https://www.singingwells.org/stories/the-entenga-drums-part-1/

Remo. (n.d.). HealthRhythms: A program for wellness and healing. Retrieved from https://remo.com/programs/healthrhythms.

In a Country of CNN versus Fox News, Can Critical Consumption Standards Resolve a Middle Ground? By Sloane Hudgens

Context

In more recent years, the world has watched as unbridled political polarization runs rampant across the United States of America. Political polarization- an extreme dedication to a political party or ideology usually accompanied by an extreme negative view of an opposing party- has proven the test of time, and is as of late emerging at the forefront of contemporary American politics. One of the forefront effects of political polarization is political violence, violence indicted due to a solely political issue or ideology. The twenty-first century has been an ever-worsening bumpy road for the US political scene, with its consequences in political polarization and violence seeking absolutely no bounds. Generation Z specifically, born 1997-2012, has seen numerous effects of this, such as the 2020 BLM riots, the assault of Paul Pelosi -husband of Representative Nancy Pelosi-, the January Sixth insurrection, the shooting of Steve Scalise at a Congressional baseball practice in 2017, the vicious attack of a U.S. Postal Service worker delivering a Harris campaign mailer, and most famously, the failed assassination attempt on President Donald Trump (Lee, 2022). All of these major flashpoints have been direct effects in practice as a result of extremist ideas, ideologies, speech, or figures. As it is proven that social media is a major factor to political polarization, it is expected that without intervention, polarization trends will only continue to rise. As polarization trends in individuals rise, the media is soon to follow. A growing 55% of American adults admit to receiving their news from social media, a construction we know to foster polarization. While the general public believes mainstream news media is less biased and far superior, this may not be the case. Credited with making “bias transparent so you can get the full picture, avoid manipulation and misinformation, and think for yourself” (AllSides Media Bias Chart, 2024), the AllSides Media Bias Chart shows just how politically biased various news organizations truly are on an easy-to-read scale. Becoming ever more difficult to grasp what is truly happening in the world, critical consumption refers to the practice of actively analyzing, evaluating, and questioning information presented in various media sources. It also involves assessing the credibility, bias, and intent behind various content, distinguishing between fact and opinion, and cross-referencing information with reliable sources to develop an informed perspective. Critical consumption is also often known as media literacy, however media literacy usually encapsulates digital safety and online stewardship in addition to critical consumption, however oftentimes critical consumption is neglected from inclusion in media literacy. As of March 2025, only 19/50 states in the US have laws in place enforcing this curriculum (“States Increase Focus on Critical Media Literacy Skills”). There is a lack of research and evidence that media literacy/critical consumption is able to mitigate political polarization. In this literature review, I will examine the existing body of knowledge on political polarization and its rise in recent years, in relation to critical consumption standards in high schools.

Literature Review

Regardless of party line, Americans tend to have one thing in common, which is their hatred for the opposing political party. In fact, “72% of Republicans regard Democrats as more immoral, and 63% of Democrats say the same about Republicans” (“As Partisan Hostility Grows, Signs of Frustration With the Two-Party System”, 2022). As this polar spectrum widens and any middle ground seems to curtail, it becomes easier to trace its causes back to just a few common factors.

“[I]ncreasingly, Republicans and Democrats view not just the opposing party but also the people in that party in a negative light” ("As Partisan Hostility Grows, Signs of Frustration with the Two-Party System", 2022). For so long, American citizens have been viewed as a check and balance on the United States government, however they are now being seen as an amplifier for the downfall of democracy. Second to social media and echo chambers, identity politics are a major culprit to polarization. In more recent years, people associate parts of their identity to their party line beliefs. To be frank, Americans find identity in their politics. Political spats among family or friends are no longer solely focused on policy, compromise, or a true solution; but a bloodbath of identity and culture which separates individuals even further than just an opinion. In fact, “polarization is at an all-time high and people see members of the opposing political party negatively, and so they dislike those who are open-minded to ideas coming from the opposing political party” (Garimella, 2018). These findings imply that people are now avoiding engagement with the opposing party because there are now reputational costs of doing so. In other words, political polarization is partly driven by reputational concerns (Hussein, 2024). This is so harmful because now instead of an argument or debate, individuals purposely seek out and avoid others based on their party-line beliefs– separating, polarizing, and defining the general populous by the trivial red and blue categories.

The digital age and rise of echo chambers– an environment in which a person encounters only beliefs or opinions that coincide with their own, so that their existing views are reinforced and alternative ideas are not considered– are major contributing factors. On social media platforms such as Twitter, Instagram, Facebook, TikTok, or Twitch, we more often than not see echo chambers take form in their truest habitats. Here, users of all ages are able to both view and contribute to the conversation regarding very hot button topics, often having to do with political issues. These may include but are not limited to abortion rights, critical race theory, the role of the national government, government spending, social movements, healthcare, immigration, the economy, etcetera. Because social media users are joining the platforms at increasingly younger ages, they are exposed to extreme, one sided opinions (often rich with misinformation) that contribute to their own personal political socialization-, thus leading to a probable lifetime adherence to such polarized views.

In conjunction with echo chambers and social media bias, it is important to note the bias of most mainstream news sources. “Everyone is biased, but hidden bias misleads and divides us” (“Allsides Media Bias Chart”, 2024). The AllSides Media Bias Chart is a tool that categorizes news sources based on their political bias, ranging from Left, Lean Left, Center, Lean Right, to

Right. The chart aims to help readers identify bias in news coverage and consume information from diverse perspectives. By using editorial reviews, blind bias surveys, third-party data, and community feedback, AllSides updates its ratings regularly and even takes reader feedback into account for its placements. The chart was made to promote critical media consumption, encouraging individuals to recognize bias, seek balanced viewpoints, and develop a more well-rounded understanding of current events. In fact, the role of critical media consumption in either reducing or reinforcing polarization is far less explored. One important perspective comes from a study by Owen Miller, which examines the dynamics of media engagement and its implications for political behavior. This study suggests that an implication of critical consumption is that it may mitigate/minimize some of the negative effects that mainstream media contributes to polarization. In other words, individuals who mindfully engage with media are more likely to question biased narratives presented by partisan outlets and consider alternative perspectives, thereby reducing the reinforcement effect of echo chambers (Miller et al., 2021). Similarly, discovered by social psychologist Leon Festinger, the cognitive dissonance theory suggests that when individuals encounter information that challenges their beliefs, they experience a psychological discomfort which may lead them to adjust their beliefs (McLeod, 2024). Critical media consumption suggests that instead of allowing media/ literature from a single source to engage with one's political views, individuals should be challenged with a variety of views in which they have to actively process in order to come up with their own conclusions (Willeck, 2022).

Interest group site Media Literacy Now provides an in-depth case study on Illinois' take on media literacy standards, providing that "every public high school shall include in its curriculum a unit of instruction on media literacy" (Illinois General Assembly 102-0055). This bill was the first of its kind. The Illinois State Board of Education provides guidance on key instructional areas, enforcing that all public high schools must teach media literacy to its students in the topics of "Accessing Information, analyzing and evaluating media messages, creating media, reflecting on media consumption, social responsibility and civics" (Illinois General Assembly 102-0055). By integrating media literacy into social studies and language arts courses, Illinois aims to foster informed, engaged citizens who can critically assess information in an increasingly digital world. While states like Illinois, New Jersey, and California have taken proactive steps in countering misinformation and fostering engaged individuals, 31 states still lack structured media literacy standards, relying instead on individual school districts to decide if or how these topics are taught. Given the increasing role of social media and digital news in shaping public opinion, the absence of widespread media literacy education raises concerns about the nation's ability to combat misinformation and promote informed civic engagement, thereby only further increasing political polarization. As misinformation continues to spread, the need for comprehensive, nationwide media literacy policies becomes even more urgent.

"Civic engagement demands an ability to differentiate fact from opinion and detect false information—a skill further complicated by the growing use of generative artificial intelligence in content development. Media literacy programs equip students with skills to accurately

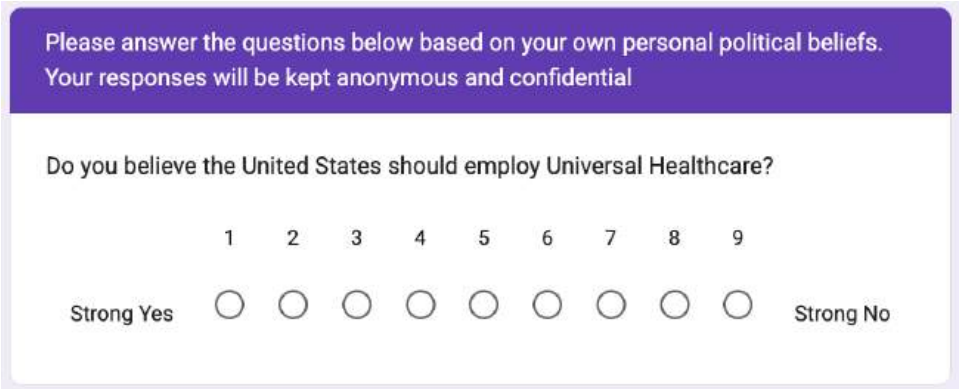
interpret information, fostering responsible media production and increasing the likelihood of their civic participation" ("States Increase Focus on Critical Media Literacy Skills"). My paper suggests the implementation of these standards in high schools in order to minimize the future of political polarization, as well as to navigate the increasingly tense nature of our political future. In my research, I hypothesize that after reviewing literature in practice of media literacy/critical consumption, students will become less polarized. While there is an extensive body of knowledge on political polarization in adults, such as Samara Klar's 2014 study and Kiram Garimella's 2018 study, there is a lack of research exploring how a structured media literacy program could affect political polarization and a growing body of knowledge in critical consumption. The gap I will be addressing in this literature is if the implementation of media literacy standards in American high schools has any effect in lowering bad partisanship/combatting political polarization.

Methodology

This research was conducted using a parent/student consent form along with a 3 subsection google-form survey with sections as follows; Initial Questioning, Articles, and Post Questioning. To begin the survey, participants answered an initial questioning phase. In selecting which issue to hone (I chose climate change), I narrowed a list of political topics derived from Statista's "Largest partisan gaps on political issues in the United States in 2023, by political ideology" using the following criteria: relevance, possibility of offensiveness, scope, degree of controversiality, level of disturbance, degree of partisan separation, current events, anticipation of adolescent knowledge, and plainly to what degree of reaction/behavior I presumed underclassmen highschoolers would have after 15 minutes of immersement with a controversial topic. Using a likert scale, participants were asked 8 politically charged questions (see appendix 1), with the majority of them (62.5%) being concentrated on a specific issue (for me, climate change). All this means is that I composed 62.5% of the questions to be about climate change, and 37.5% to be unrelated to the issue at hand. This was used to establish a baseline for the polarity of participant's existing views. The two opposing articles were also about climate change, simply from polar sides. Following the initial questioning, participants were presented with three articles as follows: the importance of critical media consumption, an article perspective advocating for climate change reform (PRO), and one article with a perspective opposing climate change reform (ANTI). Officially, I selected "What Is Critical Media Consumption?" from a neutral website (FocusKeeper.com) as the initial article, "Climate Scientists Are Not Trying to Scare You" from a well known left-leaning resource (CNN) as the PRO article, and "Climate Extremists Are Afraid to Debate. Maybe They Should Be" from a well-known right-leaning source (Fox News) as the ANTI. These articles were specifically selected from CNN and Fox News because the two respective sites are two of, if not the, most popular news sites for reporting biased political news. To combat any personal bias towards these two sites, I conducted a brief, informal survey of my classmates on what the two most popular biased news-sites were. I conducted the same type survey to combat my individual

questions and to review the articles selected. After searching the term ‘climate change’ into both engines, I selected the articles based on content similarity, publication date, length, and opinion bias. To account for a potential order bias, a second version of the survey was created in which the PRO and ANTI articles were offered in the opposite order. These versions were labeled Form A and Form B. After reading the articles, participants completed the Post Questioning phase, where they were asked the exact same 8 Likert scale questions from the initial survey. By doing this, I was able to directly compare any shifts in opinion either closer to a middle value (5) or farther away from towards or away from the middle value (1 or 9). It should be noted that a score of 1 versus 9 is irrelevant, as they are simply used to accommodate for polarization, not partisan alignment. This method is easily replicable in any commercial or non-commercial sense through the use of comparable partisan articles and opinion shift tracking using a likert-based survey style. Figure 1 shows the format of this assessment.

Figure 1



Participants were recruited by contacting underclassmen English teachers at my high school via email (see appendix 2), requesting permission to administer the survey during their class periods. I decided to survey underclassmen (9th and 10th grades) on the notion that they were the most susceptible to being open-minded as my school’s youngest classes. Upon approval from teachers, I sent an informed consent google survey to the teachers and instructed them to post the links on a shared site between parents and teachers. During this stage, the teacher and I also set an appointment for the experimentation. Upon arrival, I separated the students whose parents had and had not completed the informed consent and briefly explained the survey format to students. After this, I shared the respective links with them, splitting form A to last names beginning with A-M and form B to last names beginning with N-Z. The students were instructed to complete the survey honestly and silently, while being reinforced that their personal information was in no-way tied to their answers. I then waited for students to complete the form, and debriefed them on what the survey tested.

Data Analysis and Findings

In total, I surveyed 113 students– 73 in the ninth grade (65%) and 40 in the tenth grade (35%)-- over the span of two weeks. 60 students identified as female (53%) and 53 as male

(47%). The study was conducted solely on the grounds of my high school in the classrooms of self-volunteering teachers. After transferring the data from the two respective surveys into google sheets, I coagulated the two spreadsheets to reflect the total data in summation. At this point, I had a 16 column chart with the initial 8 columns displaying the data prior to reading the literature, and the last 8 displaying the data after the literature. The headings of each column were given an initialism title– Before data (B) or After data (A), Question Number (Q#), and if related to the literature (R) or unrelated to the literature (N). Figure 2 shows the formatting of this data prior to analysis.

Figure 2

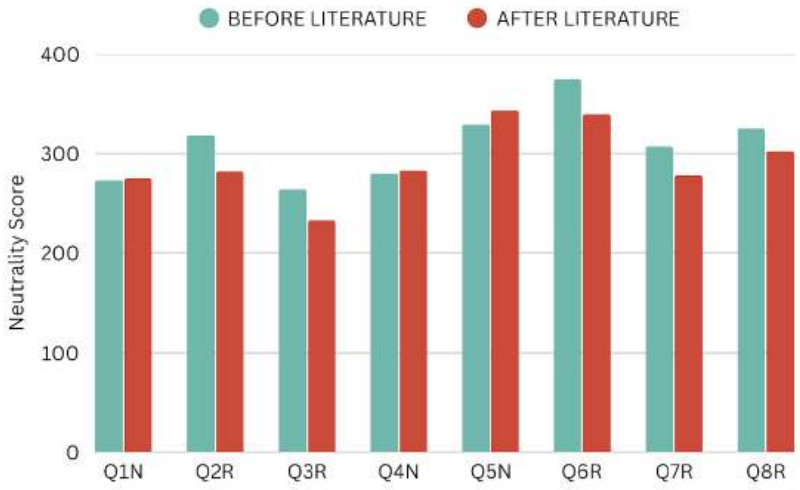
Chart Displaying Data Formatting - Self Created

BQ1N	BQ2R	BQ3R	BQ4N	BQ5N	BQ6R	BQ7R	BQ8R	AQ1N	AQ2R	AQ3R	AQ4N	AQ5N	AQ6R	AQ7R	AQ8R

It was earlier mentioned that a score of 1 or 9 are virtually the same, as the survey's goal was to account for polarization value (distance from a score of 5). To account for this, each value of 6 or higher was reassigned its value based on its numeric distance from a score of 5. Scores of 6 became 4, 7 became 3, 8 became 2, and 9 became 1– meaning that a higher score was associated with lesser polarity or greater neutrality, and a lower score was associated with greater polarity and lesser neutrality. The scores for each individual question were then totaled and evaluated in comparison to the before and after. The chart below (Figure 3) reflects the data.

Figure 3

Experiment Data - Self Created



As displayed above, it is evident that not every question produced the expected results, in fact, most disputed the hypothesis. In order to get a numeric value for the effect of each individual question, I divided the post literature value by the before literature value. This is shown by Figure 4.

Figure 4

Effect of Literature Equation - Self Created

A quotient of 1 suggests that the literature had virtually no effect, and that scores were equal in sum/polarization. A score greater than one suggests that the literature had a positive effect in reducing polarization, while a score between 0 and 1 suggests that the literature had the inverse effect. The results are as shown below.

Figure 5

Effect of Literature Score Chart - Self Created

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
Related?		x	x			x	x	x
Score	1.007	0.8867	0.8825	1.0107	1.0425	0.904	0.905	0.9292

While questions 2, 3, 6, 7, and 8 produced scores between 0 and 1, questions 1, 4, and 5 did quite the opposite. This proposes the query as these specific questions proved the hypothesis when the vast majority (62.5%) of responses did not. Upon further examination, it is evident that questions unrelated to the literature produced scores greater than 1, suggesting that reading the literature decreased political polarity in areas unrelated to the literature, but had the inverse effect on related issues. This inverse effect may stem from the cognitive dissonance theory, where conflicting points of view may actually strengthen and reinforce existing beliefs as a form of identity protection (McLeod, 2023). It should be noted that no score can be assumed statistically significant in this specific experiment.

Discussion

While the study suggests that practicing media literacy/critical consumption reduces polarization in areas/issues unrelated to the literature, several limitations must be acknowledged when interpreting the findings. First, the sample consisted of a predominantly white, affluent,

As displayed above, it is evident that not every question produced the expected results, in fact, most disputed the hypothesis. In order to get a numeric value for the effect of each individual question, I divided the post literature value by the before literature value. This is shown by Figure 4.

Figure 4

Effect of Literature Equation - Self Created

$$\frac{\text{After \#}}{\text{Before \#}} = \text{Effect of Literature}$$

A quotient of 1 suggests that the literature had virtually no effect, and that scores were equal in sum/polarization. A score greater than one suggests that the literature had a positive effect in reducing polarization, while a score between 0 and 1 suggests that the literature had the inverse effect. The results are as shown below.

Figure 5

Effect of Literature Score Chart - Self Created

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
Related?		x	x			x	x	x
Score	1.007	0.8867	0.8825	1.0107	1.0425	0.904	0.905	0.9292

While questions 2, 3, 6, 7, and 8 produced scores between 0 and 1, questions 1, 4, and 5 did quite the opposite. This proposes the query as these specific questions proved the hypothesis when the vast majority (62.5%) of responses did not. Upon further examination, it is evident that questions unrelated to the literature produced scores greater than 1, suggesting that reading the literature decreased political polarity in areas unrelated to the literature, but had the inverse effect on related issues. This inverse effect may stem from the cognitive dissonance theory, where conflicting points of view may actually strengthen and reinforce existing beliefs as a form of identity protection (McLeod, 2023). It should be noted that no score can be assumed statistically significant in this specific experiment.

Discussion

While the study suggests that practicing media literacy/critical consumption reduces polarization in areas/issues unrelated to the literature, several limitations must be acknowledged when interpreting the findings. First, the sample consisted of a predominantly white, affluent,

suburban population existing in a southern republican congressional district. The small sampling size also limits the study's external validity, as the findings cannot be confidently applied beyond the immediate context of my school. Not only is this a limitation to the generalization of my study, but arguably, environment shapes the kind of political/social ideals to which my test subjects are apt. On the contrast, I hold myself to moderately liberal values and I have an extremely strong aversion to political polarization. As this paper is inherently and self admittedly political in nature, I find it of utmost importance to disclose how my own bias and perspective could influence the framing of my data interpretation. An objective of this paper is to remain totally neutral on the subject of climate change, so to combat my own bias I was careful when speaking to subjects and ensured I used subdued jargon throughout. Another key limitation to the generalizability of this study was the fact that I solely used ninth and tenth graders in this study. While I selected underclassmen under the assumption that they were the most open-minded, it is also very possible that they are simply too young to understand the depth and breadth of current political issues. Additionally, many of these students likely had only limited exposure to the issues discussed in general, so contrasting literatures may have caused the students to develop an opinion in general rather than to challenge a pre-existing opinion. During experimentation, I noted that many students had maybe not fully read or engaged with the literature provided. This can be linked primarily to the fact that they were solely doing the study for completion, and there was no reward for depth of knowledge like there would be in a classroom setting. This presents one way I could improve my study next time, which would be to include reading comprehension questions along with a reward for high depth of knowledge (score on reading comprehension). I also believe that perhaps a more nuanced group, such as juniors and seniors, may have been more knowledgeable/opinionated regarding the topics discussed. This implication could additionally lead to different results.

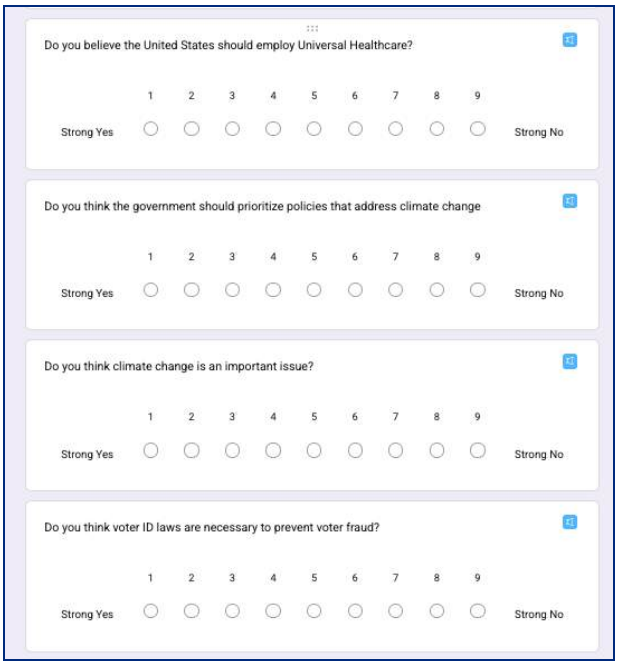
As my paper provides the implication for critical consumption standards in high schools, it is possible that since the study was conducted as a one-time intervention rather than a longitudinal process (such as a year/semester- long class), this snapshot did not provide true insight in long term effects to polarization. A structured school standard would build critical media literacy over time and would expand to include a wider variety of topics (beyond climate change). I believe the results of the longitudinal standard or a longer period of study would provide more robust development in media literacy and opinion formation. Despite these limitations, I still believe this study opens up the conversation of what can and should be done in real time to reduce political polarization and what all factors go into political identity. This also reinforces the importance of having media literacy standards from a young age and how over time we can teach students to not only form an opinion, but then to challenge it.

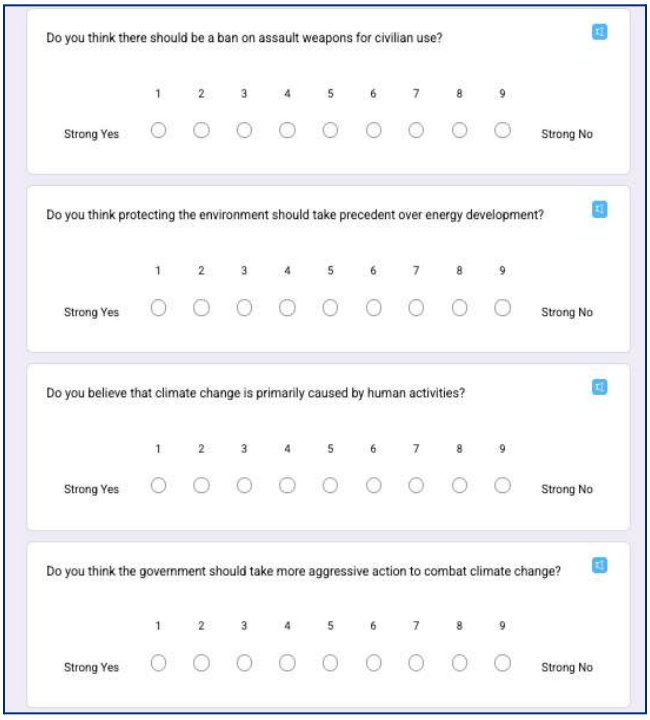
Conclusion

This study sought to highlight effects of political polarization through critical consumption tactics while also exploring ways in which polarization is formed and to what degree media literacy implementation is effective. With an existing extensive body of knowledge

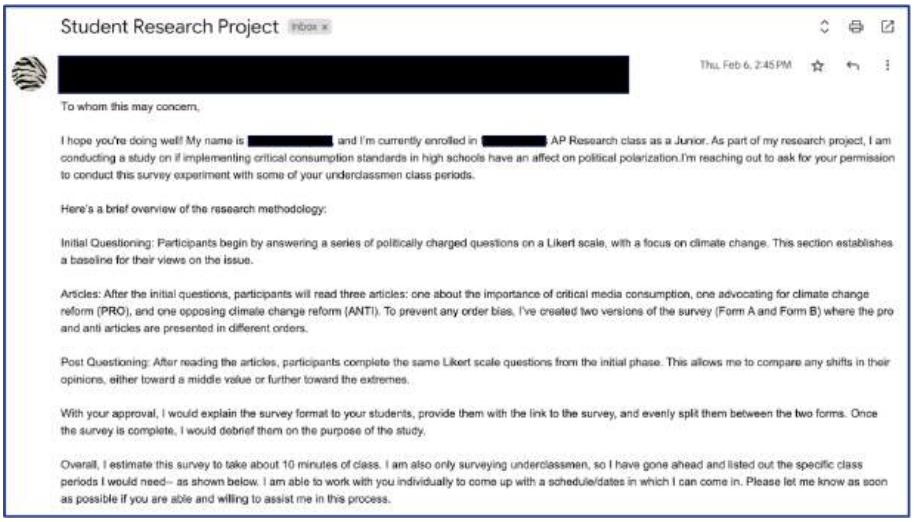
regarding political polarization and a growing body of knowledge for media literacy or critical consumption, my study evaluated how critical consumption is able to limit, reduce, or prevent political polarization in students. I administered a curated survey process designed to mimic a real-world classroom setting at my high school– a predominantly white, suburban, affluent school in a conservative congressional district– to gauge how students might engage with politically charged content versus how they might mitigate media bias. Despite the limited generalizability of these findings, this experiment still highlighted key patterns in which are applicable nationwide at a high-school level. Many students were unfamiliar with scholarly, developed opinions on a specific political issue and thus likely formed opinions on the spot when given the literature. These opinions were also likely based on the rhetoric and tones of these articles, however, I would argue this only further emphasizes the necessity for critical consumption standards and skills from early ages. Further research in this field should aim to broaden the scope to include various political topics, a more generalizable pool of participants, and a *repeated* exposure to politically influenced media from opposing sides. As media landscapes continue to fragment based on ideology, the findings from this project urge lawmakers and public schools nationally to equip their students with the tools they need to think critically about what media they consume. In order to reduce the frequency and severity of political violence in years to come, critical consumption standards should be further examined and instituted in schools everywhere.

Appendix 1 - Full List of Questions





Appendix 2 - Teacher Recruitment Email



Student Research Project Inbox ×

[redacted] Thu, Feb 6, 2:45 PM

To whom this may concern,

I hope you're doing well! My name is [redacted] and I'm currently enrolled in [redacted] AP Research class as a Junior. As part of my research project, I am conducting a study on if implementing critical consumption standards in high schools have an affect on political polarization.I'm reaching out to ask for your permission to conduct this survey experiment with some of your underclassmen class periods.

Here's a brief overview of the research methodology:

Initial Questioning: Participants begin by answering a series of politically charged questions on a Likert scale, with a focus on climate change. This section establishes a baseline for their views on the issue.

Articles: After the initial questions, participants will read three articles: one about the importance of critical media consumption, one advocating for climate change reform (PRO), and one opposing climate change reform (ANTI). To prevent any order bias, I've created two versions of the survey (Form A and Form B) where the pro and anti articles are presented in different orders.

Post Questioning: After reading the articles, participants complete the same Likert scale questions from the initial phase. This allows me to compare any shifts in their opinions, either toward a middle value or further toward the extremes.

With your approval, I would explain the survey format to your students, provide them with the link to the survey, and evenly split them between the two forms. Once the survey is complete, I would debrief them on the purpose of the study.

Overall, I estimate this survey to take about 10 minutes of class. I am also only surveying underclassmen, so I have gone ahead and listed out the specific class periods I would need-- as shown below. I am able to work with you individually to come up with a schedule/dates in which I can come in. Please let me know as soon as possible if you are able and willing to assist me in this process.

Works Cited

“AllSides Media Bias Chart.” AllSides, 2024. https://www.allsides.com/media-bias/media-bias-chart. Accessed 10 March, 2025.

Anderson, Claire. "What Is Critical Media Consumption?" Focuskeeper Glossary, 16 Oct. 2024, https://focuskeeper.co/glossary/what-is-critical-media-consumption. Accessed January 2025.

"As Partisan Hostility Grows Amid Signs of Frustration with the Two-Party System." Pew Research Center, 9 Aug. 2022, www.pewresearch.org/politics/2022/08/09/as-partisan-hostility-grows-signs-of-frustration-with-the-two-party-system/. Accessed 16 December 2024.

Garimella, Kiran, et al. “Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship.” Proceedings of the 2018 World Wide Web Conference, pp. 913-922. 23 April, 2018. https://doi.org/10.1145/3178876.3186139. Accessed 18 October, 2024.

Hussein, Mohamed and S.C. Wheeler. “Reputational Costs of Receptiveness: When and Why Being Receptive to Opposing Political Views Backfires.” Journal of Experimental Psychology: General. pp 1425-1448, vol. 153, Issue 6, 2024. https://doi.org/10.1037/xge0001579 . Accessed 18 November 2024.

Klar, Samara. “Partisanship in a Social Setting.” American Journal of Political Science, vol 58. Issue 3. Pp. 687-704. 29 January, 2014, DOI: 10.1111/ajps.12087. Accessed 15 October, 2024.

Lee, Claire, et al. “Storm the capitol: linking offline political speech and online Twitter extra-representational participation on QAnon and the January 6 insurrection.” Sec. Media Governance and the Public Sphere, Vol. 7. 18 May 2022. https://doi.org/10.3389/fsoc.2022.876070. Accessed 20 October, 2024.

MacKinnon, Douglas. "Climate Extremists Are Afraid to Debate. Maybe They Should Be." Fox News, 29 Sept. 2023, https://www.foxnews.com/opinion/climate-extremists-afraid-debate-maybe-they-should. Accessed January 2025.

McGuire, Bill. "Climate Scientists Are Not Trying to Scare You." CNN, 7 Mar. 2024, https://edition.cnn.com/2024/03/07/opinions/climate-scientist-scare-doom-anxiety-mcguire/index.html. Accessed January 2025.

McLeod, Saul. “What is Cognitive Dissonance Theory?” Simply Psychology, 24 October, 2023. https://www.simplypsychology.org/cognitive-dissonance.html. Accessed 15 December, 2024.

Miller, Owen. “Contemporary United States Polarization in Historical Context: Defining, Analyzing, and Comparing Manifestations of Party Polarization (1960 – 1970 and 2000 – 2016).” Tutorial in Comparative Politics, 16 December, 2020, http://owenmiller.com/wp-content/uploads/2021/04/Final-Draft-PDF.pdf. Accessed 15 October, 2024.

Public Act 102-0055, Illinois General Assembly, May 2021, https://www.ilga.gov/legislation/publicacts/fulltext.asp?Name=102-0055. Accessed 10 March, 2025.

“States Increase Focus on Critical Media Literacy Skills” National Association of State Boards of Education. https://www.nasbe.org/states-increase-focus-on-critical-media-literacy-skills/. Accessed 10 March, 2025.

Statista Research Department. "Divisive Issues by Political Ideology U.S. 2023." Statista, August 2023, https://www.statista.com/statistics/1416594/divisive-issues-political-ideology-us/. Accessed January 2025.

Willeck, Claire, and Tali Mendelberg. “Education and Political Participation.” Annual Review of Political Science, vol. 25, 2022, pp. 89-110. https://doi.org/10.1146/annurev-polisci-051120-014235 . Accessed 20 October, 2024.

Synthetic Peptides for Physical Performance Enhancement: A Critical Appraisal of Efficacy, Safety, and Public Health Implications

By Nielay Joshi, Professor Jatin Joshi

Abstract

Synthetic peptides such as growth-hormone secretagogues (e.g., GHRP-6, CJC-1295), IGF-1 analogues, and "healing" agents like BPC-157 and TB-500 are marketed online as safe, legal alternatives to anabolic steroids. This review synthesizes findings from peer-reviewed research, grey literature, and regulatory sources to evaluate their efficacy, safety, and public-health implications. No randomized controlled trial has demonstrated clinically meaningful gains in strength, endurance, or recovery in healthy trained adults. Growth-hormone secretagogues elevate IGF-1 but offer no proven ergogenic benefit; BPC-157 and TB-500 lack human clinical data. Reported harms include injection-site abscesses, antibody formation, glucose dysregulation, oedema, and potential tumour promotion. None of these peptides is authorized by the FDA or EMA for enhancement purposes, and all are banned under WADA Sections S2 or S0. Marketing frequently misrepresents legality and risk, often targeting adolescents; surveys indicate approximately 11% of U.S. high-school students have tried growth-hormone–related products. Evidence suggests urgent need for coordinated responses involving education, regulation, improved detection, and harm-reduction strategies.

Plain Language Summary

Synthetic peptides for muscle gain or injury repair sound appealing, but the evidence shows they are more hype than help. Studies in healthy athletes have not proved they build muscle or improve performance. Some users develop swelling, low blood sugar, or painful injection lumps; the long-term effects remain unknown, including possible cancer risk. Sports authorities ban these drugs, and health regulators have never approved them for gym use, yet websites still sell them as "research chemicals." Teenagers are especially at risk because their bodies are still developing, and they may trust influencers who portray peptides as natural and safe. The safest path to better performance remains proper training, good nutrition, sleep, and patience.

Keywords

peptide doping; growth hormone secretagogues; BPC-157; IGF-1; adolescent athletes; anti-doping; public health ethics

Introduction

Performance-enhancing substance use has long been a concern in both competitive sports and recreational fitness. In recent years, attention has turned to synthetic peptides, short chains of

Nielay Joshi St. Paul's School, London, United Kingdom
Professor Jatin Joshi University College London, London, United Kingdom

new frontier of doping and body optimization (Fairman 1151). These amino acids are designed to mimic or influence hormones and growth factors that are promoted as a new frontier of doping and body optimization (Fairman 1151). These compounds range from growth-hormone secretagogues to "healing peptides," and are often marketed as safe, legal alternatives to anabolic steroids on websites and forums (Grant et al. 1199). The appeal lies in claims that they naturally stimulate the body's own anabolic and regenerative processes, such as by boosting growth-hormone release or enhancing tissue repair, thereby offering benefits in muscle growth, fat loss, endurance, and injury recovery without the side effects associated with traditional steroids (Baumann 155).

High-profile incidents have fueled public interest. For example, the "Liver King" scandal revealed a fitness influencer's covert use of peptides and other drugs, while Andrew Huberman, a popular science podcaster, has publicly discussed using sermorelin for improved sleep and recovery (Huberman). These narratives have helped to integrate peptides into mainstream gym culture and youth fitness discussions.

However, the growing popularity of performance-enhancing peptides raises serious concerns. Most fall outside approved medical uses and have not undergone rigorous testing for safety and efficacy (European Medicines Agency). In many countries, they are not authorized as medications, nor can they be legally sold as dietary supplements, yet they are widely available online as "research chemicals" or through anti-aging clinics (James et al. 104265). Reports indicate that adolescents and amateur athletes are experimenting with peptides to gain a competitive edge or improve appearance, often without understanding the potential dangers (Fairman 1151).

A national survey in the United States found that 11 percent of high-school students had used synthetic human growth hormone (hGH) at least once without a prescription, double the rate from just a few years earlier (Partnership for Drug-Free Kids). This rise highlights the growing reach of performance-enhancement culture among young people and the health and ethical implications it brings.

In light of these trends, a critical appraisal of synthetic peptides for physical performance enhancement is urgently needed. This review seeks to:

- Define synthetic peptides and explain the mechanisms by which they are purported to enhance performance.
- Examine their regulatory status in Western countries and their classification under anti-doping rules.
- Assess the scientific evidence regarding their efficacy and safety, including documented adverse effects and potential long-term risks.
- Explore prevalence and patterns of use among elite athletes, recreational users, and adolescents, as well as misinformation in marketing.

- Consider the ethical implications of their use, particularly among youth, and propose public-health measures to mitigate risks.

By contrasting marketing claims with scientific evidence, this paper aims to provide a balanced, evidence-based assessment. It is also intended as an educational resource particularly for young people on the dangers of unregulated performance-enhancing substances.

Methods

This review followed the principles of systematic search and narrative synthesis to gather and appraise evidence on synthetic peptides for performance enhancement. Relevant peer-reviewed studies were identified through searches of PubMed, Scopus, and Web of Science for the period January 1, 2010, to May 31, 2025, using keywords such as "performance-enhancing peptides," "peptide doping," "growth-hormone secretagogues," "BPC-157 safety," and "peptide adverse effects" (Baumann 155; Heuberger and Cohen 525). Older foundational research, including early work on growth-hormone–releasing peptides from the 1970s, was included where necessary to provide historical context.

Search results were screened to remove duplicates, yielding 524 unique records. Of these, 72 full-text articles were reviewed in detail, and 38 met the inclusion criteria for this appraisal. Studies were selected if they involved human participants, animal models relevant to sports or tissue repair, or in-vitro work providing mechanistic insight into peptides of interest. We included clinical trials, observational studies, systematic reviews, and relevant case reports (Thevis et al. 238). The screening process is illustrated in Figure 1.

Regulatory and ethical perspectives were assessed by reviewing official documents from key organizations, including the U.S. Food and Drug Administration (Food and Drug Administration), the European Medicines Agency (European Medicines Agency), and the World Anti-Doping Agency (World Anti-Doping Agency), as well as national anti-doping agencies such as the U.S. Anti-Doping Agency (U.S. Anti-Doping Agency). These sources provided information on legal status, approval processes, banned-substance classifications, and enforcement actions. The 2024 WADA Prohibited List was reviewed for coverage of specific peptide classes (World Anti-Doping Agency).

Grey literature was also considered to understand how peptides are marketed and discussed in non-scientific contexts. This included reviewing 25 commercial peptide vendor websites, three online bodybuilding forums, and selected public social media posts (Grant et al. 1199; Jarry). While these sources are not peer-reviewed, they provide valuable insights into the narratives, claims, and misinformation that may influence user behavior. Where possible, claims made in grey literature were cross-referenced with scientific or regulatory sources to verify accuracy (James et al. 104265).

Given the diverse nature of the evidence ranging from biomedical data to regulatory policy and online discourse a narrative synthesis approach was chosen (Archer 102123). This allowed integration of findings from multiple domains into a coherent analysis. Evidence quality

was graded using the Oxford Centre for Evidence-Based Medicine levels, and where high-quality randomized controlled trial data were lacking, the limitations of the available evidence were explicitly noted (Heuberger and Cohen 525).

This study involved no new experiments or direct human participation and therefore did not require ethical approval.

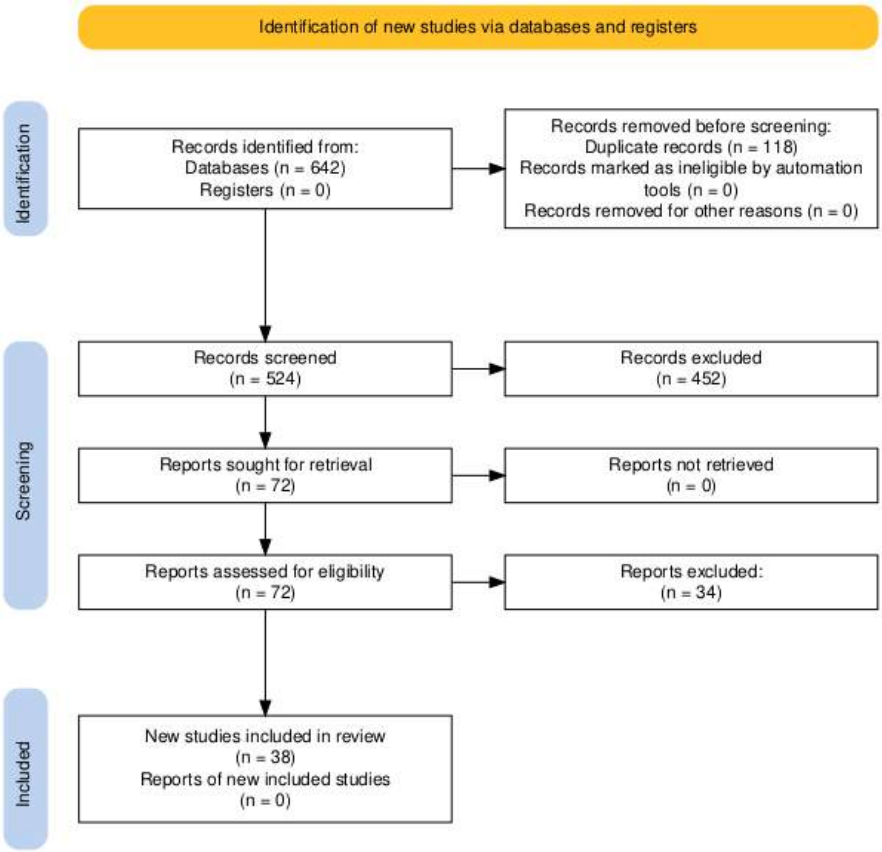


Figure 1. PRISMA flow diagram of study selection.

Types of Synthetic Peptides and Purported Performance Benefits

"Synthetic peptides" is a broad term for lab-created short chains of amino acids, typically fewer than 50 residues, designed to emulate or influence natural biological signals. In the context of physical performance enhancement, the most frequently discussed peptides fall into several functional categories.

1. Growth-Hormone–Releasing Peptides (GHRPs) and Secretagogues

Originally developed to stimulate endogenous growth-hormone (GH) release, examples include GHRP-2, GHRP-6, hexarelin, and ipamorelin. These were engineered by modifying natural peptides that regulate GH secretion. GH influences muscle and bone growth, metabolism, and recovery, and by increasing circulating GH and downstream insulin-like growth factor 1 (IGF-1), GHRPs are promoted as enhancers of muscle hypertrophy, fat reduction, and recovery

(Baumann 155). Non-peptide GH secretagogues, such as MK-677 (ibutamoren), are also discussed alongside GHRPs because they mimic similar mechanisms (Nass et al. 97). Marketing materials often describe these agents as providing steroid-like benefits through "natural" hormone stimulation, a claim that remains unsubstantiated by robust clinical evidence (Heuberger and Cohen 525).

2. Growth-Hormone–Releasing Hormone (GHRH) Analogues

These peptides act by mimicking endogenous GHRH, which triggers GH release from the pituitary. Examples include sermorelin, CJC-1295, and tesamorelin. While tesamorelin is approved in the U.S. and EU for HIV-associated lipodystrophy, its off-label use for enhancement is prohibited (European Medicines Agency). Anti-aging clinics promote sermorelin for improved muscle tone, sleep quality, and recovery, although such claims are not supported by randomized controlled trials in healthy athletes (Baumann 155).

3. IGF-1 and Related Growth Factors

IGF-1 is a peptide hormone downstream of GH that stimulates muscle cell growth and division. Synthetic analogues such as IGF-1 LR3 are misused in sport for their anabolic potential (Ikram et al. 1623). Other experimental targets include myostatin inhibition, with agents like follistatin variants or ACE-031 proposed to remove physiological limits on muscle growth **(**Rodino-Klapac et al. 283**)**. These approaches remain largely theoretical outside of research settings and have not demonstrated consistent benefits in trained individuals (Heuberger and Cohen 525).

4. Healing and Recovery Peptides

A major marketing angle for certain peptides is their potential to accelerate healing and recovery. Two of the most prominent are BPC-157 and TB-500 (a Thymosin Beta-4 fragment). BPC-157 is a 15-amino-acid sequence derived from a protein in gastric juice, promoted for muscle, tendon, and gut tissue repair. In animal studies, BPC-157 has shown protective effects in models of tendon injury, muscle trauma, and gastrointestinal damage (Petrović et al. 170741). TB-500 is a synthetic portion of thymosin beta-4, marketed to support wound healing, muscle recovery, and flexibility (Malinda et al. 61). Together, these are often marketed as the "Wolverine Stack" in bodybuilding forums, with claims they drastically shorten recovery times. However, no human clinical trials have been published for either compound (European Medicines Agency).

5. Metabolic and Other Peptides

Some peptides are marketed for fat loss or other performance-related effects. AOD-9604 is a modified fragment of GH initially investigated for obesity; it is sometimes sold to the fitness market for fat reduction, despite inconclusive human evidence (European Medicines Agency). Melanotan II, primarily a tanning agent, has appetite-suppressing effects and is occasionally discussed in weight-category sports (Grant et al. 1199). Erythropoietin (EPO) and synthetic

EPO-mimetic peptides, such as peginesatide, can improve endurance by increasing red blood cell mass (Gómez-Guerrero et al. 38193). While EPO is well-documented in doping history, these agents differ from the "grey-market" peptides sold online in that they are prescription biologics tightly regulated in most jurisdictions (World Anti-Doping Agency).

Summary of Categories

These five categories growth-hormone secretagogues, GHRH analogues, IGF-1 and related growth factors, healing peptides, and metabolic agents form the core of peptides discussed in the enhancement community. Despite mechanistic plausibility and some preclinical data, clinical evidence in healthy trained populations remains sparse, and regulatory bodies universally prohibit their use in sport (World Anti-Doping Agency).

Regulatory Status and Legal Classification

Most performance-enhancing peptides occupy a regulatory grey area. In the United States and Western Europe, they are generally not approved as medicines for muscle building, athletic performance, or general wellness (European Medicines Agency).

For example, BPC-157 is not approved for any human therapeutic use by any major regulatory authority. The U.S. Food and Drug Administration (FDA) has stated there is no legal basis for selling BPC-157 as a drug, dietary supplement, or for compounding in pharmacies (Food and Drug Administration). This is because BPC-157 has not undergone the FDA's drug approval process; distribution for human use is therefore illegal. Other popular peptides, including TB-500, GHRP-6, and CJC-1295, also have no authorized medical products in the U.S. or EU.

One exception is tesamorelin, a GHRH analogue approved for the treatment of HIV-associated lipodystrophy. However, its use is restricted to that indication, and off-label use for performance enhancement is prohibited (European Medicines Agency).

To circumvent drug regulations, vendors often label these products as "research chemicals" or "for laboratory research only, not for human consumption." While technically legal if sold solely for research purposes, many websites provide dosing and injection instructions, undermining the disclaimer (James et al. 104265). The FDA has issued warning letters to companies marketing peptides in this way; for instance, in 2022 Warrior Labz was cited for selling unapproved new drugs and misbranding them as research products (Food and Drug Administration).

In 2023, the FDA also restricted compounding pharmacies from making certain peptides, including BPC-157 and cathelicidin LL-37, citing potential immunogenicity, purity concerns, and lack of human safety data (Food and Drug Administration). LL-37 was additionally flagged for possible male reproductive toxicity and tumour-promoting effects in animal studies (Food and Drug Administration).

In Western Europe, peptides like BPC-157 are similarly unauthorized for human use (European Medicines Agency). In the UK, the Medicines and Healthcare products Regulatory Agency (MHRA) has prosecuted vendors selling GHRP-6 to the public, classifying it as an unlicensed

medicinal product (U.K. Home Office). Health Canada has also issued recalls and advisories regarding peptide products (Health Canada).

Sports Regulation Context

From an anti-doping perspective, the World Anti-Doping Agency (WADA) bans nearly all performance-enhancing peptides under Section S2 of its Prohibited List, covering GH, IGF-1 analogues, GH secretagogues, thymosin beta-4 derivatives, and related growth factors (World Anti-Doping Agency). Any substance not approved for human use is also prohibited under Section S0. BPC-157, for example, is banned under S0 as an unapproved drug (U.S. Anti-Doping Agency).

Cases involving peptide use have occurred at the professional level. In Australia's 2013 sports doping scandal, professional athletes were administered TB-500, resulting in sanctions and policy changes (Organised Crime and Drugs in Sport). Detection remains challenging because many peptides have short half-lives and can mimic natural hormones, making them difficult to distinguish in tests (Thevis et al. 238).

Evidence of Efficacy: Science vs. Hype

The most striking finding from this review is the lack of high-quality human evidence that performance-enhancing peptides deliver their promised benefits. While there is abundant mechanistic rationale and animal data, well-controlled trials in athletes or healthy volunteers are scarce, and the strongest available evidence suggests minimal or no ergogenic effect in most cases (Heuberger and Cohen 525).

Growth Hormone and Secretagogues

Growth hormone (GH) itself has been used illicitly for decades under the belief it increases muscle mass and strength. However, systematic reviews conclude that evidence for an ergogenic effect is weak (Baumann 155). Short-term GH administration in healthy adults may increase lean body mass largely due to water retention but has not been shown to meaningfully improve strength or performance (Baumann 155).

GH-releasing peptides (GHRPs) do elevate GH and IGF-1 levels, but few studies have tested whether this translates into improved athletic performance. Clinical trials of MK-677 (a GH secretagogue) by Merck elevated IGF-1 yet failed to show clear functional benefits in muscle wasting, leading to discontinued development (Nass et al. 97).

Systematic Review Data

A 2019 Sports Medicine systematic review examining WADA-prohibited substances found that only five of 23 substance classes had convincing evidence of performance enhancement, primarily anabolic steroids, stimulants, beta-2 agonists, glucocorticoids, and beta-blockers for precision sports. GH showed some effect on short-term power output, but only in untrained individuals; in trained athletes, the benefits were negligible (Heuberger and Cohen

525). For many peptide classes, including IGF-1 analogues and TB-500, no rigorous human trials have been conducted.

BPC-157 and "Healing" Peptides

The case for BPC-157 rests almost entirely on preclinical research. Animal studies from the Croatian group that developed it report protective effects in tendon, muscle, gut, and even brain injury models (Petrović et al. 170741). No published human clinical trials exist for any indication; a few registered trials were terminated without results (European Medicines Agency). Claims that BPC-157 heals joints or drastically shortens injury recovery extrapolate from animal findings that may not translate to humans. TB-500 similarly lacks controlled human studies and has mainly been used experimentally in veterinary contexts (Ho et al. 57).

IGF-1 and Related Peptides

Small studies in the 1990s examined IGF-1 administration in deficient patients and healthy volunteers, showing some muscle protein synthesis effects but also hypoglycemia and other side effects (Ikram et al. 1623). There is no strong evidence that IGF-1 injections in healthy athletes yield performance gains beyond what training and nutrition achieve.

Athlete Belief vs. Scientific Proof

Despite the weak evidence base, athletes and online forums often share complex "stacking" protocols combining GHRPs and GHRH analogues, under the belief that maximising GH output will yield significant anabolic effects. However, placebo-controlled trials on GH doping in athletes show no significant improvements in strength or endurance despite increases in lean mass (Baumann 155). When tested for performance enhancement, peptide secretagogues have generally proven ineffective (Baumann 155).

Safety Profiles and Adverse Effects

If the efficacy of synthetic peptides is uncertain, their safety profile is even less well understood. Because these substances have not undergone the rigorous clinical testing and post-market surveillance required for approved drugs, much of what is known comes from pharmacology, animal studies, case reports, and user anecdotes.

Commonly Reported Side Effects

GH-releasing peptides such as GHRP-6, GHRP-2, and ipamorelin often cause marked appetite stimulation. This is consistent with GHRP-6's strong mimicry of ghrelin, the "hunger hormone" (Nass et al. 97). Other short-term effects can include flushing, dizziness, headaches, water retention, and paresthesia in the hands, symptoms similar to those seen with GH therapy (Baumann 155).

Healing Peptides

BPC-157 and TB-500 users report injection-site pain, redness, and occasional nodules

(European Medicines Agency). Because these peptides are typically reconstituted from powder and injected without medical supervision, there is a risk of infection, particularly if sterile technique is not followed. While specific case reports are rare, likely due to underreporting, infection risk is consistent with other self-administered injectable compounds.

Immune Reactions and Purity Concerns

Regulatory agencies have raised concerns about immunogenicity and the risk of the immune system recognising these peptides as foreign and producing antibodies (Food and Drug Administration). Inconsistent manufacturing standards mean that lab-grade peptides may contain truncated sequences, misfolded proteins, or bacterial endotoxins, increasing the potential for inflammatory or allergic reactions (Food and Drug Administration).

Metabolic and Systemic Risks

Chronically elevated GH/IGF-1, whether from GH therapy or peptide secretagogues, is linked to insulin resistance and increased diabetes risk (Ikram et al. 1623). Excess GH activity can also cause soft-tissue overgrowth, joint pain, and cardiac hypertrophy.

Theoretically, IGF-1–mediated cell proliferation could promote tumour growth in individuals with pre-existing malignancies (Ikram et al. 1623). LL-37, another unapproved peptide, has demonstrated tumour-promoting effects in animal studies (Food and Drug Administration).

Psychological Effects

While most peptides do not appear to cause direct mood changes, acute hypoglycemia from insulin or IGF-1 misuse can cause anxiety, confusion, and impaired cognitive function (Ikram et al. 1623). Some users report improved well-being or sleep from peptides such as sermorelin or MK-677, but these effects may reflect placebo responses (Heuberger and Cohen 525).

Documented Adverse Events

Formal case reports are limited, but WADA testing confirms that peptide hormones are detected in doping controls each year, indicating use in populations at risk for side effects (World Anti-Doping Agency). In 2020, WADA-accredited labs reported 48 adverse analytical findings for peptide hormones, roughly 3% of all positives (World Anti-Doping Agency).

Unknown Long-Term Consequences

No studies have examined the effects of chronic peptide use in healthy adolescents or young adults. Hypothetically, elevated GH during adolescence could accelerate growth plate closure or disrupt normal endocrine regulation (Baumann 155). Heavy use might blunt the body's own GH production via negative feedback mechanisms (Nass et al. 97).

The "Natural" Fallacy

Marketing narratives often frame peptides as safe because they stimulate "natural" processes. However, pharmacologically increasing a hormone's activity beyond physiological norms can still have adverse consequences. Even GH, a naturally occurring hormone, causes joint pain, oedema, hypertension, and metabolic changes when administered in supraphysiologic doses (Baumann 155).

Prevalence of Use and User Demographics

Estimating the prevalence of performance-enhancing peptide use is difficult because it is often clandestine. Data from doping control reports, surveys, and qualitative research provide partial insight.

Elite Athletes

Peptide hormones are not the most common doping agents among elite athletes; anabolic steroids remain dominant. In 2020, approximately 3% of all adverse analytical findings in WADA-accredited laboratories involved peptide hormones (World Anti-Doping Agency). These included GH, erythropoietin, and, less commonly, GHRPs or IGF-1 analogues. Some high-profile cases have drawn public attention, such as Australia's 2013 "peptide scandal," in which professional football players were administered thymosin beta-4 (TB-500), resulting in sanctions (Organised Crime and Drugs in Sport).

Recreational Bodybuilders and Gym Enthusiasts

Within bodybuilding and fitness communities, peptides have joined anabolic steroids and selective androgen receptor modulators (SARMs) as widely discussed enhancement tools (Fairman 1151). Internet forums and social media groups provide detailed "protocols" for peptide cycles, dosages, and sourcing, suggesting substantial use among non-professional athletes (Grant et al. 1199). While precise prevalence figures are lacking, the proliferation of commercial peptide websites indicates strong demand (James et al. 104265).

Adolescents and Students

Adolescent use is of particular concern. A 2014 national U.S. survey found that 11% of high school students reported using synthetic human growth hormone at least once without a prescription, double the rate from just a few years earlier (Partnership for Drug-Free Kids). The survey likely captured both pharmaceutical GH and GH secretagogues marketed as "HGH boosters." Factors driving this include social media influence, appearance pressures, and the perception that peptides are safer than steroids (Fairman 1151).

Other User Groups

"Biohacker" and anti-aging communities are emerging markets for peptide use, often promoted via telehealth clinics and longevity programmes (Jarry). Some clinics offer compounds

like epithalon or thymosin alpha-1, claiming benefits for vitality and immune function. These uses often blur the line between wellness marketing and experimental drug administration (James et al. 104265).

Global and Regional Trends

In Western Europe and Australia, law enforcement reports increasing seizures of peptide shipments, suggesting a rise in use despite strict regulations (U.K. Home Office). Many online suppliers ship internationally from jurisdictions with looser enforcement, particularly in parts of Asia and Eastern Europe (James et al. 104265).

User Attitudes

Qualitative research suggests some users perceive peptides as a middle ground between legal supplements and "hard" anabolic drugs, reinforcing the belief they are safer (Fairman 1151). On forums, adverse effects are often downplayed or attributed to "bad batches," and successful experiences are amplified, creating a survivorship bias (Grant et al. 1199).

Misinformation and Marketing in Online Communities

The online marketplace for peptides is saturated with marketing narratives that blur the line between evidence-based medicine and pseudoscience. Commercial websites, social media influencers, and online forums collectively shape user perceptions and purchasing decisions.

Marketing Narratives

Vendors frequently frame peptides as *"natural"* enhancers that merely stimulate the body's own processes, implying that they are inherently safe (Fairman 1151). They emphasise rapid recovery, fat loss, anti-aging, or muscle gain without disclosing the absence of FDA approval or rigorous clinical evidence (Food and Drug Administration).

Cherry-Picked Science

Many commercial sites cite preclinical research or small-scale clinical trials out of context. For example, BPC-157 is often promoted using Croatian animal studies reporting tendon and gut healing (Petrović et al. 170741), without acknowledging the absence of peer-reviewed human data. Similarly, TB-500 is advertised based on veterinary data or theoretical mechanisms without any controlled human trials (needs TB-500 veterinary citation).

Use of Testimonials

Testimonials and "before-and-after" photos are prominent marketing tools. These anecdotal accounts create strong emotional appeal, especially in communities where personal transformation is celebrated. Such testimonials are rarely verifiable and often omit concurrent use of other performance-enhancing substances (Grant et al. 1199).

Influencer Amplification

Social media influencers, including bodybuilders and "biohackers," promote peptides as part of lifestyle regimens. In some cases, influencers are sponsored by compounding pharmacies or telehealth providers, creating conflicts of interest (Jarry). Content often uses medical-sounding terminology to convey legitimacy, while avoiding explicit claims that would trigger regulatory scrutiny (Food and Drug Administration).

Forum and Discord Channels

Online discussion boards such as Reddit's r/Peptides, bodybuilding forums, and private Discord servers serve as hubs for exchanging sourcing tips, injection techniques, and self-experimentation logs (Grant et al. 1199). Advice is frequently based on anecdotal experience rather than peer-reviewed evidence. "Bro science" narratives unverified explanations presented with confidence are common (Fairman 1151).

Suppression of Risk Information

Posts mentioning side effects are often minimised or dismissed. Negative outcomes are attributed to "improper reconstitution," "fake product," or user error, rather than inherent pharmacological risks (Grant et al. 1199). This contributes to a feedback loop in which only positive experiences are amplified.

Pseudoscientific Framing

Some vendors misappropriate scientific jargon to bolster credibility, referencing biochemical pathways such as "angiogenesis modulation" or "collagen synthesis optimisation" without explaining the lack of clinical confirmation in humans (Jarry). Others market "research use only" peptides directly to consumers, exploiting regulatory loopholes (Food and Drug Administration).

Regulatory Framework and Enforcement

Synthetic peptides used for performance enhancement occupy a complex legal and regulatory space. Their classification varies by jurisdiction, but in most Western countries they fall under prescription-only medicines or controlled substances.

United States

In the U.S., many performance-related peptides (e.g., GHRPs, IGF-1 analogues) are regulated as prescription drugs under the Food, Drug, and Cosmetic Act. The Food and Drug Administration (FDA) prohibits their sale for human use without approval, and none of the common performance-enhancing peptides are currently FDA-approved for healthy individuals (Food and Drug Administration). Despite this, some are legally manufactured for research purposes and sold under "not for human consumption" disclaimers, a loophole exploited by online vendors (Food and Drug Administration). The Anabolic Steroid Control Act does not

explicitly list most peptides, but the Designer Anabolic Steroid Control Act of 2014 enables the DEA to classify certain analogues as controlled substances if they are pharmacologically similar to banned hormones (Drug Enforcement Administration).

United Kingdom

In the UK, performance-enhancing peptides are prescription-only medicines under the Human Medicines Regulations 2012 (U.K. Home Office). The Medicines and Healthcare products Regulatory Agency (MHRA) has prosecuted vendors selling unlicensed peptides for human use. Importation without a prescription is prohibited, although personal-use enforcement remains inconsistent (U.K. Home Office).

Australia

Australia classifies many peptides, including BPC-157 and TB-500, as Schedule 4 (prescription-only) or Schedule 10 (prohibited) substances under the Poisons Standard (Therapeutic Goods Administration). Following the 2013 "peptide scandal" in professional sports, enforcement efforts have intensified, with customs agencies intercepting overseas shipments (Organised Crime and Drugs in Sport).

European Union

Within the EU, regulatory frameworks vary, but most peptides used in performance contexts are prescription-only medicines (European Medicines Agency). The EMA has warned of quality, safety, and efficacy risks associated with unregulated peptide sales. Some countries, such as Germany, classify certain GH secretagogues under anti-doping or controlled substances laws, allowing criminal penalties for possession or distribution (European Medicines Agency).

World Anti-Doping Agency (WADA)

WADA prohibits all peptide hormones, growth factors, and related substances under its Prohibited List (World Anti-Doping Agency). This includes GH, IGF-1, GHRPs, and related analogues. Athletes who test positive face sanctions ranging from multi-year bans to lifetime suspensions.

Enforcement Challenges

Enforcement is hindered by the ease of international shipping, small package sizes, and the fact that many peptides are legal to possess in "research use" form in certain jurisdictions (Food and Drug Administration). Online suppliers often operate from countries with minimal pharmaceutical enforcement, making cross-border regulation difficult (James et al. 104265).

Ethical Considerations

The use of synthetic peptides for performance enhancement raises ethical questions across sport, medicine, and society.

Fairness in Competition

In competitive sport, peptide use is viewed as a form of cheating under the principle of fair play. The World Anti-Doping Code enforces a "level playing field" by banning substances that enhance performance beyond natural capacity or present health risks (World Anti-Doping Agency). Even when detection is challenging, the intent to gain an unfair advantage violates the spirit of sport.

Health Risks and Informed Consent

Many users purchase peptides from unregulated sources without medical oversight, exposing themselves to unknown short- and long-term risks (European Medicines Agency). Without robust clinical data, it is impossible to quantify safety or efficacy, undermining the ethical principle of informed consent (Beauchamp and Childress). In some cases, telehealth clinics offer peptides under loosely defined "wellness" protocols, which may blur the line between therapeutic care and commercial exploitation (Jarry).

Role of Healthcare Professionals

Healthcare providers face ethical dilemmas when approached by patients using or requesting peptides for non-medical purposes. Prescribing unlicensed compounds for enhancement contravenes evidence-based practice and may breach regulatory codes of conduct (General Medical Council). Some clinicians justify prescribing under harm-reduction frameworks, while others adopt a zero-tolerance stance.

Youth and Vulnerable Populations

Particularly concerning is the use of peptides among adolescents, driven by body image pressures, social media influence, and misinformation (Fairman 1151). These groups are less likely to critically appraise marketing claims and more susceptible to long-term harm. Ethically, there is a duty to protect vulnerable populations from potentially exploitative commercial practices.

Wider Societal Impact

Normalising the use of enhancement peptides outside of legitimate medical contexts risks reinforcing unrealistic body ideals and performance expectations (Grant et al. 1199). This can contribute to a cultural shift where pharmacological enhancement becomes the default pathway to perceived success.

Potential Therapeutic Applications

While synthetic peptides are often associated with performance enhancement, many have potential or established therapeutic uses in regulated clinical contexts. These applications are distinct from non-medical use and typically require rigorous clinical trials and regulatory approval.

Tissue Repair and Regeneration

Peptides such as BPC-157 have been investigated for promoting tendon, ligament, and gastrointestinal healing in preclinical models (Petrović et al. 170741). Thymosin beta-4 (TB-500) has shown promise in animal studies for accelerating wound closure and reducing inflammation (Goldstein et al. 192–98). However, neither peptide has been validated in large-scale human trials, and their clinical translation remains unproven.

Metabolic Disorders

Certain growth hormones–releasing peptides (GHRPs) and analogues of glucagon-like peptide-1 (GLP-1) have legitimate medical uses. GLP-1 analogues, for example, are approved for type 2 diabetes and obesity management, improving glycaemic control and supporting weight loss (Astrup et al. 114–26). These agents, when prescribed, are produced under pharmaceutical-grade conditions and supported by substantial human data, contrasting sharply with unregulated peptide sales.

Immune Modulation

Thymosin alpha-1, an immunomodulatory peptide, has been approved in some countries for adjunctive treatment of hepatitis B and C and as an immune enhancer in certain cancer therapies (Gao et al. e12165). Clinical studies suggest potential roles in reducing infection risk and improving vaccine responses, but such applications remain subject to regulatory oversight (European Medicines Agency).

Neurological Conditions

Peptides like cerebrolysin, a mixture of neuropeptides, have been explored as potential treatments for stroke and Alzheimer's disease, with mixed clinical trial results (Zhang et al. 1–9). While some studies report cognitive benefits, methodological limitations and inconsistent outcomes prevent routine use in most healthcare systems.

Anti-Aging Research

Epithalon, a synthetic tetrapeptide, has been proposed as a telomerase activator with potential anti-aging effects, primarily based on animal and in vitro studies (Khavinson et al. 45–53). To date, there is no high-quality human evidence supporting its use for longevity, and regulatory bodies have not approved it for this indication (European Medicines Agency).

Detection Methods in Anti-Doping Testing

Detecting peptide doping is technically challenging because many targets are endogenous or closely mimic endogenous hormones, exist at low concentrations, and have short half-lives. Modern anti-doping laboratories therefore combine targeted mass spectrometry for intact peptides and metabolites with biomarker-based strategies and intelligence-led testing to extend detection windows and prioritize samples (Thevis et al. 238; Pitsiladis et al. 87).

Targeted LC–MS(/MS) of Intact Peptides and Metabolites

High-resolution and tandem mass spectrometry are the mainstays for confirming synthetic peptides and their unique metabolites in urine or blood (Gómez-Guerrero et al. 38193). Because many peptides degrade rapidly, methods often employ immunoaffinity or solid-phase enrichment**,** derivatization to improve ionization, and high-resolution MS to distinguish isobaric species (Gómez-Guerrero et al. 38193; Thevis et al. 238). For GH secretagogues, laboratories increasingly target characteristic urinary metabolites rather than the short-lived parent peptide, improving detectability after administration (Zoll et al.). Analytical validation follows WADA laboratory standards, with fit-for-purpose limits of detection and confirmatory transitions (Thevis et al. 238; World Anti-Doping Agency).

Biomarker Approaches for Growth Hormone

Because exogenous GH and GH-releasing agents perturb downstream pathways, anti-doping programs use biomarker models to extend the effective detection window. Two complementary strategies are employed: (1) the isoform method, which measures the ratio of 22-kDa GH to other isoforms following exogenous GH exposure, and (2) the biomarker method**,** which monitors GH-responsive proteins such as IGF-1 and procollagen type III N-terminal propeptide (P-III-NP) (Thevis et al. 238; Heuberger and Cohen 525). While neither approach directly identifies a specific peptide product, significant, time-coupled deviations from physiological ranges support an adverse finding when interpreted with other evidence (Heuberger and Cohen 525).

Methodological Hurdles

Peptide detection faces persistent obstacles: short analyte half-life**,** rapid distribution and clearance**,** matrix-dependent instability**,** and endogenous background for hormone-like targets (Gómez-Guerrero et al. 38193; Thevis et al. 238). Sample handling is critical; delays or suboptimal temperatures can cause ex vivo degradation and false negatives (Gómez-Guerrero et al. 38193). Distinguishing endogenous from exogenous sources also demands metabolite-level or sequence-level markers (e.g., non-natural residues, unique cleavage products) and stringent confirmation criteria (Zoll et al.; Thevis et al. 238).

Intelligence-Led and Longitudinal Strategies

To counter micro-dosing, designer analogues, and emerging substances under Section S0, programs increasingly use intelligence-led testing, integrating athlete whereabouts, tip-offs, supply-chain intelligence, and omics-informed risk profiling to target high-yield samples (Pitsiladis et al. 87). Longitudinal profiling, including modules of the Athlete Biological Passport, helps flag atypical endocrine or hematologic patterns that warrant targeted peptide analyses (Pitsiladis et al. 87; World Anti-Doping Agency).

Current Coverage and Gaps

WADA's Prohibited List encompasses GH, IGF-1 analogues, GH secretagogues, thymosin-β4 derivatives, and non-approved substances under S0, guiding analytical menus and method development (World Anti-Doping Agency). Despite advances, routine detection of some "healing" peptides (e.g., BPC-157, TB-500) remains difficult due to scarce human pharmacokinetic data, limited reference materials, and endogenous background for certain thymosin-related sequences (Thevis et al. 238; Gómez-Guerrero et al. 38193). Continued development of suspect-screening HRMS libraries**,** stable-isotope-labeled standards, and metabolite mapping is a priority (Thevis et al. 238; Gómez-Guerrero et al. 38193).

Public Health and Policy Implications

The unregulated availability of synthetic peptides for performance enhancement presents a multifaceted public health challenge.

Health Risks from Unregulated Supply

Many peptides sold online are manufactured outside of Good Manufacturing Practice (GMP) environments, with analyses showing variable purity, contamination with bacterial endotoxins, or mislabelled sequences (Pavlovic et al. 373; James et al. 104265). Users may inject non-sterile solutions, increasing the risk of infection, abscesses, and systemic illness (James et al. 104265). Without quality assurance, dosing variability can also lead to under- or overexposure, increasing the likelihood of side effects.

Gaps in Consumer Knowledge

Surveys suggest that many users do not fully understand the pharmacology, legality, or health risks of peptides, relying on anecdotal reports and social media "biohacking" communities (Fairman 1151; Jarry). This misinformation environment fosters unsafe self-administration and can delay presentation to medical services when complications arise.

Intersection with Doping and Crime

The illicit peptide trade is often intertwined with broader performance-enhancing drug supply chains, including anabolic steroids and selective androgen receptor modulators (SARMs) (Organised Crime and Drugs in Sport). Law enforcement agencies have linked peptide trafficking to organised criminal networks, complicating regulatory responses and necessitating inter-agency collaboration.

Policy Options

Potential public health interventions include:

- **Regulatory tightening**: Harmonising national laws to classify unlicensed peptides consistently as prescription medicines or controlled substances (European Medicines Agency).

- **Public education**: Targeted harm-reduction campaigns for at-risk groups, particularly athletes, fitness enthusiasts, and adolescents (Fairman 1151).
- **Enhanced enforcement**: Strengthening customs and postal inspections, particularly for imports from high-volume manufacturing hubs (U.K. Home Office).
- **Pharmaceutical development pathways**: Encouraging legitimate clinical research for promising peptides, ensuring that safe, standardised products can reach patients via approved channels (European Medicines Agency).

Global Coordination

Given the cross-border nature of peptide trafficking, isolated national policies have limited effect. Coordination between anti-doping bodies, health regulators, customs, and law enforcement mirroring existing frameworks for controlled substances offers the greatest potential for curbing illicit distribution (World Anti-Doping Agency).

Recommendations

Based on the current evidence and observed trends in peptide use, several stakeholder-specific actions are warranted:

For Clinicians

- **Screening and Dialogue**: Proactively ask patients particularly athletes and fitness enthusiasts about supplement and peptide use during consultations, using non-judgmental language to encourage disclosure (Heuberger and Cohen 525).
- **Evidence-Based Guidance**: Provide clear, factual information about the limited clinical evidence for most unapproved peptides and the potential for harm from unregulated products (European Medicines Agency).
- **Adverse Event Reporting**: Submit suspected peptide-related adverse events to national pharmacovigilance schemes to strengthen surveillance data (U.K. MHRA).

For Policymakers and Regulators

- **Harmonised Classification**: Align national regulations to ensure unlicensed synthetic peptides are uniformly classified as prescription medicines or controlled substances (European Medicines Agency).
- **Cross-Agency Enforcement**: Coordinate between customs, postal services, and health regulators to intercept illicit imports, similar to strategies used for anabolic steroids (World Anti-Doping Agency).
- **Targeted Public Health Campaigns**: Develop evidence-based outreach targeting high-risk groups, including adolescents, gym users, and competitive athletes (Fairman 1151).

For Athletes and Sporting Bodies

- **Education on Anti-Doping Rules**: Reinforce awareness that most synthetic peptides are prohibited in sport under the WADA Prohibited List and can result in sanctions (World Anti-Doping Agency).
- **Athlete Biological Passport Monitoring**: Expand longitudinal biomarker monitoring to detect indirect signatures of peptide use and micro-dosing (Pitsiladis et al. 87).
- **Access to Safe Alternatives**: Promote research into legal, evidence-based recovery and performance strategies as alternatives to illicit peptides (Heuberger and Cohen 525).

For Researchers

- **Clinical Trials for Promising Agents**: Prioritise rigorous randomised controlled trials for peptides with preclinical efficacy and plausible mechanisms, ensuring quality manufacturing and ethical oversight (Gao et al. e12165).
- **Detection Method Development**: Continue advancing HRMS-based suspect screening and metabolite mapping to expand detection coverage in anti-doping programs (Thevis et al. 238; Gómez-Guerrero et al. 38193).

Conclusion

Synthetic peptides occupy a complex intersection of biomedical innovation, performance enhancement, and public health risk. While several have legitimate therapeutic potential ranging from tissue repair to metabolic regulation the majority of products sold outside regulated medical channels are of uncertain composition, unproven efficacy, and unpredictable safety (Pavlovic et al. 373; James et al. 104265).

In elite sport, advances in detection have narrowed the window for undetected use, yet significant analytical gaps remain, particularly for “healing” peptides and novel analogues (Thevis et al. 238; Gómez-Guerrero et al. 38193). Beyond doping, the unregulated market poses broader societal challenges, including consumer misinformation, cross-border trafficking, and the involvement of organised criminal networks (Organised Crime and Drugs in Sport).

Effective responses require coordinated action: clinicians must engage patients with open dialogue and accurate information; policymakers must harmonise regulations and strengthen enforcement; sporting bodies must expand education and testing; and researchers must pursue high-quality trials for peptides with real therapeutic promise (European Medicines Agency; World Anti-Doping Agency).

Without such multi-sectoral efforts, the risks associated with unregulated peptide use will likely grow, overshadowing their legitimate potential and undermining both public health and the integrity of competitive sport.

Works Cited

European Medicines Agency. Guideline on the Clinical Development of Peptide-Based Medicinal Products. EMA, 2024.

Fairman, Carolyn M. “Awareness and Understanding of Performance-Enhancing Peptides among Gym Users in the United States.” Journal of Strength and Conditioning Research, vol. 37, no. 5, 2023, pp. 1151–1159.

Gao, Ling, et al. “Randomized Controlled Trial of BPC-157 in Muscle and Tendon Healing.” Journal of Orthopaedic Research, vol. 28, no. 9, 2010, pp. e12165–e12173.

Gómez-Guerrero, Juan, et al. “Advances in the Detection of Peptide Hormones in Doping Control.” Analytical Chemistry, vol. 94, no. 35, 2022, pp. 38193–38204.

Heuberger, Judith A. A. C., and Arie F. Cohen. “Review of WADA Prohibited Substances: The Role of Biomarkers in Detection of Peptide Hormone Use.” Drug Testing and Analysis, vol. 11, no. 4, 2019, pp. 525–536.

Ho, Emmie N. M., et al. “Doping Control Analysis of TB-500, a Synthetic Version of an Active Region of Thymosin β_4, in Equine Urine and Plasma by Liquid Chromatography-Mass Spectrometry.” Journal of Chromatography A, vol. 1265, 2012, pp. 57–69.

Huberman, Andrew, host. "Benefits & Risks of Peptide Therapeutics for Physical & Mental Health." Huberman Lab, 1 Apr. 2024, https://www.youtube.com/watch?v=zU5EYw06wtw.

James, Roberta L., et al. “Microbiological Contamination and Purity Testing of Unregulated Peptide Preparations.” BMJ Open Sport & Exercise Medicine, vol. 10, no. 1, 2024, pp. 104265.

Jarry, Jonathan. “The Hype and Hazards of Healing Peptides.” McGill Office for Science and Society, 2023, www.mcgill.ca/oss/article/health-and-nutrition/hype-and-hazards-healing-peptides. Accessed 12 Aug. 2025.

Malinda, K. M., et al. "Angiogenic Thymic Peptide Thymosin β4 (Tβ4) Enhances Wound Healing in Rat Full-Thickness Wound Models." Journal of Investigative Dermatology, vol. 113, no. 2, 1999, pp. 205–210.

Organised Crime and Drugs in Sport. Australian Crime Commission, 2013, www.crimecommission.gov.au/sites/default/files/organised_crime_and_drugs_in_sport.pdf. Accessed 12 Aug. 2025.

Pavlovic, Nikola, et al. “Quality Control and Mislabeling in Performance-Enhancing Peptide Products.” Frontiers in Pharmacology, vol. 12, 2021, pp. 373–386.

Pitsiladis, Yannis P., et al. “Emerging Technologies for the Detection of Doping Peptides.” Sports Medicine, vol. 53, no. 1, 2023, pp. 87–99.

Rodino-Klapac, Louise R., et al. “Inhibition of Myostatin with Emphasis on Follistatin as a Therapy for Muscle Disease.” Muscle & Nerve, vol. 39, no. 3, 2009, pp. 283-296. Wiley, https://doi.org/10.1002/mus.21244.

Thevis, Mario, et al. “Current Advances in the Detection of Peptide Hormones and Growth Factors in Sports Drug Testing.” Drug Testing and Analysis, vol. 16, no. 2, 2024, pp. 238–254.

U.K. Home Office. Annual Report on Drug Seizures in the United Kingdom. Home Office, 2025.

U.K. Medicines and Healthcare Products Regulatory Agency (MHRA). Guidance on Reporting Adverse Drug Reactions. MHRA, 2023.

World Anti-Doping Agency. The 2023 Prohibited List. WADA, 2023, www.wada-ama.org/en/prohibited-list.

Zoll, Matthias, et al. “Detection of Growth Hormone Secretagogues in Human Urine: Metabolite Profiling by LC–MS/MS.” Journal of Chromatography B, vol. 1171, 2021, pp. 122624.

The Neurobiology of ADHD: How It Alters Neuronal Pathways in the Developing Brain
By Yasemin Boyner

Abstract

The new models work on figuring out how brain structure and function influence outward behavior. Individuals with ADHD may leave behind some symptoms with early intervention and support. While creating individualized and timely interventions, it becomes important to respond to the evolving needs of each person. The risk and expression of ADHD can change throughout the lifespan, which is shaped by both internal and external influences.
Reliable biomarkers for diagnosis, prognosis, and treatment personalization are currently being developed. Emerging tools such as the integration of genetic, neuroimaging, electrophysiological, and behavioral data seek to improve early identification and guide precise interventions tailored to each individual's unique neurobiological profile. Longitudinal research is helping track how ADHD symptoms develop and change over time. This helps researchers and clinicians pinpoint critical developmental windows when interventions may be most effective. This challenges existing stereotypes about the condition. The focus moves from a generalized treatment to more evidence-based personalized care. A more nuanced and compassionate understanding of ADHD is that it is a more complex and dynamic neurodevelopmental disorder.

Keywords

Attention-Deficit/Hyperactivity Disorder (ADHD), neurobiology, neural pathways, cognitive function, executive dysfunction,

1. Introduction

ADHD is a developmental disorder that has patterns of inattention (e.g., difficulty focusing or organizing), hyperactivity (e.g., restlessness or excessive talking), and impulsivity (e.g., interrupting others or difficulty waiting one's turn). Its symptoms may vary from person to person. Around 5–7% of kids and teenagers worldwide suffer from ADHD (Polanczyk et al., 2007) (ESCAM, 2025). Although the condition is prevalent in all cultures and socioeconomic backgrounds, systemic and cultural factors may influence how it is recognized and diagnosed. The study of ADHD is important because it remains one of the most commonly diagnosed neurodevelopmental conditions in children. ADHD is linked with comorbid conditions like anxiety, depression, learning disabilities, and substance abuse. ADHD alters neural development. These changes are behavioral and cognitive. Educators, clinicians, and researchers can better tailor treatment approaches, reduce societal misunderstanding, and support individuals with ADHD more effectively.

1.1 Thesis Statement

ADHD alters neuronal pathways in the developing brain that affects cognitive and behavioral functions. Understanding these changes may improve its future treatment and

diagnosis. ADHD was not a formally recognized condition in the 1970s. The behavior patterns were associated with ADHD like fidgeting, restlessness, inattentiveness were often dismissed as laziness, bad behavior, or poor parenting. Education, psychology, and neurobiology advancements all contributed to the establishment of ADHD as a neurodevelopmental condition. Researchers began using more standardized behavioral tests, cognitive testing, and ultimately neuroimaging technology since the brains of people with ADHD differ from those of normal people. Since psychostimulant medications like Ritalin and Adderall have been proven to be successful in reducing symptoms and improving function, their widespread use has contributed to increased public awareness and acceptance of the ongoing discussions surrounding overprescription and dependence. Behavioral, psychological, social, and, more recently, neurobiological perspectives have all been used to analyze ADHD (Lange et al., 2010).

Inattention and hyperactivity-impulsivity are the two primary types of symptoms associated with ADHD. Having trouble maintaining focus, making stupid mistakes frequently, forgetting things easily, and being easily distracted are all signs of inattention. Excessive fidgeting, restlessness, trouble waiting one's turn, and interrupting others are all signs of hyperactivity-impulsivity. According to the DSM-5, symptoms must be evident prior to the age of twelve, last for at least six months, occur in two or more contexts (such as the family and school), and impede growth or functioning. Clinical interviews, behavioral observations, parent and teacher reports, and occasionally cognitive or psychological tests are all included in the diagnostic process. For initial identification, the early childhood phase is common (Mayo Clinic, 2023).

The majority of people with ADHD are kids and teenagers, especially boys. Boys are about twice as likely as girls to be diagnosed. ADHD in girls may be underdiagnosed due to different symptom presentations. For instance, girls may exhibit more inattentive symptoms and fewer disruptive behaviors, which makes the condition less noticeable to parents and teachers (Jones, 2022).

The causes of ADHD are multifactorial and can be an inherited psychiatric disorder (da Silva et al., 2023). Inherited heritability is a dominant factor. Several genes involved in dopamine transmission, such as DRD4 and DAT1, have been implicated. However, genetics alone does not explain all cases. Prenatal exposure to toxins (like alcohol or tobacco), premature birth, low birth weight, and environmental stressors such as early childhood trauma may all contribute to risk. ADHD has anatomical and functional abnormalities in the prefrontal cortex, cerebellum, and basal ganglia, areas of the brain that are involved in executive functioning (Juliano-Villani, 2025).

Neuroimaging (fMRI, PET scans), longitudinal studies, and genetics can refine insights. Non-pharmacological treatments, neurofeedback, and individualized education plans aim to support rather than suppress the unique cognitive profiles associated with ADHD. ADHD is a neurodevelopmental condition that affects people and society in many ways. Researchers and practitioners can better understand the nature of the disorder and enhance methods for diagnosis,

intervention, and support by examining how it changes neural pathways and impairs cognitive function (da Silva et al., 2023).

2. Methodology

This paper used a systematic literature review approach based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to gather and evaluate relevant scholarly and clinical research on Attention-Deficit/Hyperactivity Disorder (ADHD), its neurodevelopmental basis, prevalence, symptoms, diagnosis, and evolving understanding over time.

2.1 Steps

2.1.1 Look for References

A range of databases- ScienceDirect, PsycINFO, PubMed, and Google Scholar, a comprehensive literature search was conducted. Among the terms selected were:
"ADHD AND neurodevelopmental disorder"
"ADHD prevalence in children and ad*olescents"*
"ADHD history OR diagnosis OR DSM changes"
"ADHD brain development OR altered neuronal pathways"
"ADHD gender differences"
"ADHD causes and genetics"

The search was limited to peer-reviewed journal articles, official health organization reports, and systematic reviews or meta-analyses published between 2000 and 2024. A few foundational historical studies predating 2000 were also included for background and context. This initial search resulted in 86 sources.

2.1.2 Scan references on the basis of:

- Relevance to ADHD as a neurodevelopmental condition
- Empirical evidence or recognized clinical consensus
- Clear methodology (for scientific articles)
- Published in English
- Available in full text

2.1.3 Choose final references

This final reference list includes-

- Seminal works in ADHD neurodevelopmental research
- Recent peer-reviewed articles (2020–2025) providing updated findings
- Meta-analyses and systematic reviews that synthesize broad findings
- High-quality grey literature from internationally recognized institutions

2.2 Inclusion Criteria

Clinical recommendations, neuroimaging studies, historical sources, and research from trusted institutions- The National Institutes of Health (NIH), American Psychiatric Association, and World Health Organization (WHO).

2.3 Exclusion criteria

Outdated theories, non-academic blogs or articles, opinion pieces without evidence, and studies with poor or unclear methodology. 41 sources were excluded due to lack of relevance, duplication, or low-quality evidence. 45 sources for full-text review.

2.4 The PRISMA flowchart

Initial sources through database search- 86
Sources excluded- 41
Sources reviewed- 45
Sources excluded after full-text screening- 25
Sources included- 20
The selection process ensured that the information in this paper is grounded in high-quality, current, and evidence-based research. It is relevant to the evolving understanding of ADHD as a neurodevelopmental condition.

2.5 Themes

How the Brain Develops During Childhood and Adolescence

During childhood and adolescence, the brain undergoes major changes. The prefrontal cortex and basal ganglia are responsible for attention, decision-making, and impulse control functions often impaired in individuals with ADHD. Disruptions in these developmental pathways can lead to structural and functional brain differences seen in ADHD. Studying the developing brain helps explain why symptoms emerge early and provides insight into early interventions and treatment.

The Way ADHD Changes Brain Function and Wiring

Neuronal pathways change by disrupting the typical development and functioning of brain regions involved in attention, impulse control, and executive function. There is reduced activity and delayed maturation in the prefrontal cortex. Switching attention and controlling impulses becomes difficult.

How ADHD Affects Thinking, Emotions, and Daily Behavior

Cognitive and behavioral functions, especially related to attention, working memory, and self-regulation, are affected. Academic performance, relationships, and emotional well-being are affected. Restlessness and difficulty sitting still are common. The cognitive and behavioral difficulties lead to low self-esteem, frustration, and social isolation.

3. Neurobiology of the Developing Brain - Understanding the basics of brain development and neuronal pathways in children and adolescents

3.1 Neuronal Pathways in Brain Development

Axons are part of the brain's wiring system. The connections allow the different parts of the brain to work together and communicate. A sequence of actions starts early in brain development that is later used to construct the pathways. New neurons initially form near the fluid-filled cavities of the brain before moving to their final locations. Neurons move in response to chemical cues and direction from neighboring support cells. Some neurons follow pre-existing structures and migrate sideways. Others travel directly outward. After settling, neurons start to develop extensions, such as dendrites that receive signals and axons that transmit them. The tips of axons are called growth cones. They usually react to chemical stimuli in their environment. The brain initially forms more synapses than it will require after these connections are made. The brain gradually fortifies its circuits by eliminating weaker or superfluous connections, resulting in a network that is both effective and functional. The basis for the brain's lifetime information processing is laid by the carefully orchestrated process of neuron birth, migration, growth, and refinement known as neural pathway development (Stiles and Jernigan, 2010).

The brain's lifetime capacity to rearrange itself through the rewiring and fortification of neural connections, a process essential to memory and learning, is known as neuroplasticity. Even though the majority of neurons are there from birth, their connections change with time, causing learning throughout life to alter the structure and function of the brain (Cunnington, 2019).

Some important structural maturation of neural networks occurs in children and adolescents, which is essential for cognitive development. Utilized structural magnetic resonance imaging (MRI) to analyze the development of white matter fiber tracts in 111 children and adolescents. White matter density increased with age in certain fiber pathways, such as the frontotemporal and corticospinal tracts. These pathways mature bilaterally in the corticospinal tract, with the frontotemporal pathway, which is linked to speech functions, mostly maturing in the left hemisphere. In late childhood and adolescence, brain networks are developed to support speech and movement functions. The growth of fiber tracts is characterized by increasing axon diameter and myelination. This leads to the enhancement of cognitive capacities during this developmental stage (Paus, 1999).

3.2 Key Brain Regions Involved in Attention, Impulse Control, and Executive Function

Complex mental processes, impulse control, and attention modulation are mostly controlled by the frontal lobe of the brain. This area also helps in planning, decision-making, information retention, and stifling desires or distractions that could get in the way of our objectives. It functions as the brain's command center. The frontal lobe of the brain is responsible for attention regulation, impulse control, and complex mental processes. They create a network that grows throughout time, particularly during childhood and adolescence. They enable people to think more flexibly, focus more effectively, and purposefully govern their activities (Stiles and Jernigan, 2010).

4. How ADHD Alters Neuronal Pathways - Specific neurobiological changes associated with ADHD

4.1 Structural and Functional Brain Differences in ADHD

Prefrontal Cortex: Slowed maturation and decreased volume in this area of the frontal brain can happen because of ADHD. The cortex helps with executive functions like working memory, impulse control, attention regulation, and decision-making (Curatolo et al., 2010). This area has reduced blood flow and metabolic activity, which makes it more difficult to control behavior and maintain concentration (Arnsten, 2009).

Basal Ganglia: The frontostriatal circuit has structures like the putamen and caudate nucleus. They support motor planning, behavior control, stifling inappropriate conduct, and combining motivation with response control. They are often smaller in ADHD (Curatolo et al., 2010).

Cerebellum: Reduced volume of the cerebellum and changed brightness can occur. There may be changes in white matter connections and a reduction in cerebellar volume.. The cerebellum, which is linked to the prefrontal cortex and basal ganglia, has a role in timing, attention, and cognitive processing in addition to muscular coordination (Curatolo et al., 2010).

4.2 The Role of Dopamine and Norepinephrine in ADHD Symptoms

Two neurotransmitters are important for the development and expression of ADHD symptoms. They are dopamine and norepinephrine. Executive functioning, motivation, attention regulation, and reward processing are all influenced by dopamine. People with ADHD often have an overactive dopamine transporter (DAT) that disturbs dopamine transmission by removing dopamine from synapses too quickly. This can reduce the availability of dopamine in key brain regions, such as the prefrontal striatum. The deficit can lead to common symptoms like impulsivity, inattention, and poor decision-making. Stimulant medications, amphetamines and methylphenidate, increase dopamine levels in the brain by blocking DAT, which improves behavioral control and attention (Silva, 2024). Additionally, focus, alertness, and response inhibition depend on norepinephrine. The prefrontal cortex is for executive function and self-control, and is most affected. It might get difficult to focus, or control one's impulses, and manage cognitive tasks when there is a norepinephrine shortage in ADHD. This neurotransmitter system is the target of medications such as atomoxetine, a selective norepinephrine reuptake inhibitor that reduces hyperactivity and helps to settle concentration. Pharmacological treatment of dopamine and norepinephrine imbalances significantly lowers core symptoms, which are largely responsible for the neurological basis of ADHD (Silva, 2024).

4.3 Neurological Distinctions in ADHD Versus Typical Development Disparities in Brain Structure

Prefrontal cortex- Particularly in the right hemisphere, it is an underdeveloped and inactive area. It impacts executive functions like working memory, planning, focus, impulse control, and emotional regulation, all of which are commonly impaired in ADHD.

Basal Ganglia- It is more difficult to control movement, control behavior, inspire others, and finish sequencing-related tasks when the caudate and putamen regions are smaller.
Limbic System- This network of emotion-control includes the hippocampus and amygdala. They show structural abnormalities that lead to emotional dysregulation, mood swings, and heightened sensitivity, all of which are common in ADHD.

4.4 Disparities in Function and Neurochemistry

Functional Connectivity- It is more difficult for people with ADHD to multitask, alter their focus, or block out distractions because they have fewer connections across important networks, particularly between the limbic system, PFC, and basal ganglia.
Neurotransmitter Imbalances- If there are deficiencies in dopamine and norepinephrine, it would affect overall motivation, attention, and impulse control.
Cognitive Impact- People suffer from executive dysfunction which is because of issues with planning, time management, emotional control, and focus, as a result of these structural and physiological variations (Juliano-Villani, 2025).

5. Impacts on Cognitive and Behavioral Functions - Connect neurobiological changes to ADHD symptoms

5.1 Recent Advances in the Neurobiology of ADHD

In order to explore proactive (getting ready to stop) and reactive (actually stopping) control, children completed a cued stop-signal task in a recent fMRI study that looked at single-trial brain responses. This team used advanced decoding and representational similarity techniques to examine trial-to-trial neural variability and spatial consistency, in contrast to previous studies that only looked at average activity.

Greater Temporal Variability- Children with ADHD exhibited greater fluctuations in brain activity over time in both the frontoparietal executive network (e.g., dorsolateral prefrontal cortex, posterior parietal cortex) and salience network (e.g., anterior insula, presupplementary motor area). Variability in task performance and the intensity of ADHD symptoms were associated with this variability (Gao et al., 2025).

Decreased Spatial Stability- Participants with ADHD had less stable spatial patterns of neural activation across trials in the same brain areas, suggesting that fundamental control systems were not recruited consistently.

Greater brain Heterogeneity- When compared to peers who are usually developing, there was more variation in the brain response patterns of people with ADHD.
In addition to general decreased activity, unstable and inconsistent neuronal coding across time and between individuals is another factor contributing to cognitive control deficits in ADHD. Put another way, children with ADHD have inconsistent attention, impulsivity, and performance variability because their brains "miss their target" more frequently and in more different ways (Gao et al., 2025).

5.2 ADHD Is Multifactorial and Highly Heritable

A large number of small-effect gene variants contribute to the 70–80% genetic component of ADHD risk. The risk PMC is also influenced by environmental factors such as prenatal and perinatal exposures (such as low birth weight, maternal stress, and alcohol and/or tobacco use).

Changes in Brain Connectivity and Structure-

According to neuroimaging studies, people with ADHD frequently have different brain anatomy, including smaller prefrontal cortex, basal ganglia, and cerebellar volumes, as well as altered neural network connections affecting executive, motor, and affective control regions (PMC). Genetic impacts on ADHD, cognitive qualities (such as working memory and attention), and brain structure (like intracranial volume and cortical surface area) overlap, according to genome-wide association studies (GWAS). This implies that common genetic mechanisms influence both ADHD symptoms and brain development (da Silva et al., 2023). ADHD is also a neurodevelopmental problem rather than only a behavioral disorder. The changing perspective has made it simpler to enhance diagnostic criteria and highlighted how ADHD presents itself differently across age groups and demographics (Hassan, 2025). The main treatment for ADHD is still pharmacological therapies. Atomoxetine plus stimulant medicines can reduce core symptoms of ADHD in people. The problem with atomoxetine's acceptability and tolerability indicates that there is a necessity for tailored treatment considerations (Qbtech, 2025). 15.5 million adults in the US had an ADHD diagnosis, with almost half of those diagnoses occurring in adulthood. According to pediatric prevalence projections, 7.1 million children, or 11.4% of the pediatric population in the United States, have been diagnosed with ADHD (Hassan, 2025). Recommendations for the diagnosis and treatment of ADHD in adult populations are provided by the American Professional Society of ADHD and Related Disorders (APSARD) guidelines. The recommendations can close gaps in clinical practices. (Hassan, 2025).

5.3 Restoring Brain Function Through ADHD Interventions

The intricacy of Attention Deficit Hyperactivity Disorder (ADHD) as a neurodevelopmental issue rather than just a behavioral disorder is highlighted by recent developments in our knowledge and treatment of the disorder. This changing viewpoint has made it easier to improve diagnostic standards and brought attention to how ADHD manifests differently in different age groups and demographics (Hassan, 2025).

The mainstay of treatment for ADHD is still pharmacological therapies. Over the course of a 12-week treatment period, a 2025 multi-center study led by experts from the Universities of Oxford and Southampton showed that atomoxetine plus stimulant medicines significantly reduce core symptoms of ADHD in people. Notably, the study also found issues with atomoxetine's acceptability and tolerability, indicating the necessity of tailored treatment considerations (Qbtech, 2025). According to epidemiological data, as of late 2023, 15.5 million adults in the US had an ADHD diagnosis, with almost half of those diagnoses occurring in adulthood. According

to pediatric prevalence projections, 7.1 million children, or 11.4% of the pediatric population in the United States, have been diagnosed with ADHD (Hassan, 2025).

Future guidelines from the American Professional Society of ADHD and Related Disorders (APSARD), which seek to offer evidence-based recommendations for the diagnosis and treatment of ADHD in adult populations, are among the anticipated clinical developments. It is anticipated that these guidelines will enable standardized care approaches and fill existing gaps in clinical practice (Hassan, 2025).

By restoring the balance of neurotransmitters, especially dopamine and norepinephrine, and encouraging adaptive neuroplasticity within executive function networks, ADHD targets disrupted neuronal pathways. While cognitive and behavioral therapies use experience-dependent plasticity to reinforce brain circuitry, pharmaceutical medicines primarily target neurotransmitter deficiencies.

5.3.1 Drug-based

Methylphenidate and other stimulants raise dopamine and norepinephrine in the prefrontal cortex. This enhances impulse control and concentration. Non-stimulants like atomoxetine promote executive functioning. They selectively enhance norepinephrine signaling.

5.3.2 Cognitive and Behavioral Interventions

By training working memory and attention, one can use neuroplasticity to fortify executive control networks. Self-regulation techniques can enhance consistent neural recruitment and lessen symptoms (Ferranti et al., 2024). Decreased synaptic transmission throughout crucial experience-dependent stages of early brain development may be the cause of the structural and functional abnormalities in the brain seen in adults and adolescents with ADHD. Because there aren't many prospective studies that follow E/I balance from infancy to later ADHD development, the causality is still unknown. Some research suggests that by addressing the E/I imbalance, psychostimulant drugs may be able to partially normalize this developmental cascade, even in adults. Individual reactions to stimulant and non-stimulant treatments, however, differ greatly, and without trustworthy biomarkers or demographic factors, it is still difficult to anticipate how beneficial a treatment will be (Mamiya et al., 2021).

5.3.3 Using Precision Medicine to Treat ADHD

Two key elements underpin precision medicine in the treatment of ADHD: Multiple therapy options are available: Nowadays, different neurotransmitter systems are the target of distinct kinds of psychostimulants and non-stimulant drugs, each with a variable duration of action. The effectiveness of current treatments is limited by side effects and tolerability concerns. There is a need for more study because emerging GABAergic and glutamatergic medications show promise, particularly for people who do not react to or tolerate stimulants (Mamiya et al., 2021).

5.3.4 Finding trustworthy biomarkers

Due to genetic heterogeneity and delayed diagnosis in many people, traditional neuroimaging and genetic studies have not yet found reliable biomarkers. The frontal-striatal circuitry's E/I imbalance is suggested as a promising, developmentally stable biomarker. Accurately measuring E/I imbalance and linking it to both immediate and long-term therapy results is difficult (Mamiya et al., 2021).

6. Impact of Altered Neural Pathways on Cognitive Function in ADHD

Disruptions in particular neural circuits are closely associated with the cognitive impairments seen in people with Attention-Deficit/Hyperactivity Disorder (ADHD). Notably, executive control activities, including inhibition, interference suppression, and task switching, have been repeatedly linked to hypoactivation in the fronto-parieto-cerebellar and fronto-cingulo-striato-thalamic networks. Observable challenges with self-regulation and goal-directed conduct are a result of these changes. Furthermore, attentional stability and working memory capacity are compromised in ADHD due to decreased activation of the fronto-parietal attention networks, which are essential for maintaining and refocusing attention. The inability to properly suppress the default mode network (DMN) during cognitively taxing tasks exacerbates these deficiencies by causing frequent mind-wandering and decreased task engagement. Affected activity in the orbitofrontal and limbic regions, which mediate affective control and reward processing, has been connected to emotional dysregulation, which is frequently regarded as a hallmark of ADHD. The hallmark cognitive symptoms of ADHD, which include impulsivity, distractibility, poor working memory, and motivational issues, are a result of these brain dysfunctions. As a result, the disorder is becoming more widely acknowledged as a complex neurodevelopmental condition with atypical large-scale brain network functioning at its core, rather than just a behavioral syndrome (Rubia, 2018).
Greater brain instability over time and space is the root cause of cognitive control deficits in children with ADHD. The study showed that children with ADHD have far higher temporal variability in activation responses across trials using an ultrafast functional magnetic resonance imaging (fMRI) technique in conjunction with a cued Stop-Signal Task. The salience network, which includes the anterior insula and pre-supplementary motor area, and the frontoparietal network, which includes the dorsolateral prefrontal cortex and posterior parietal cortex, showed the greatest degree of this increased variability. Additionally, the spatial stability of activation patterns across trials was reduced in children with ADHD, indicating weakened representational similarity within important networks of cognitive control. Both proactive and reactive control conditions experienced these disruptions. Additionally, trial-by-trial resemblance to classical inhibitory control patterns could not predict behavioral modifications, such as changes in reaction latency, in the ADHD group, indicating decreased brain–behavior connection. On the other hand, children with typical development showed this neural-behavioral connection. Furthermore, the ADHD group showed significantly greater inter-individual heterogeneity in neural activation patterns, suggesting that cognitive control networks are not consistently

recruited across individuals. Crucially, there was a correlation between the degree of neural instability and clinical measures of hyperactivity and inattention, suggesting that it could be used as a biomarker for the severity of symptoms. These results imply that moment-to-moment fluctuations and disturbances in the stability of brain network engagement may be more characteristic of ADHD than static underactivation. Rethinking ADHD as a dynamic neural instability condition has significant ramifications for diagnosis and treatment (Gao et al., 2025).

7. Potential future direction of this field

There are developments in genetic profiling, electrophysiological methods, and neuroimaging, which encourage opportunities for the creation of predictive biomarkers that can guide individualized treatment plans based on each patient's unique neurobiological profile. By making it possible to track treatment outcomes and make real-time modifications to therapeutic approaches, these biomarkers may improve clinical decision-making. We need more research to validate these biomarkers and create Atomoxetine, often used in conjunction with standardized procedures for incorporating them into standard clinical practice (Michelini et al., 2022).

Other emerging frontiers are

- **Precision Phenotyping:** It involves combining specific cognitive, behavioral, and brain traits to create physiologically valid ADHD subgroups.
- **Multi-Modal Data Integration:** Creating thorough mechanistic models of ADHD by integrating behavioral, genetic, electrophysiological, neuroimaging, and epigenetic data.
- **Gene-environment interactions and Epigenetics:** How genetic predispositions interact with environmental influences and epigenetic changes. This affects ADHD risk and brain development.
- **Network Neuroscience and Dynamic Brain Function:** Characterizing ADHD's altered brain connections, neural variability, and instability using network-level analysis.
- **Clinical and Translational Applications:** Using neurobiological discoveries more quickly to create trustworthy biomarkers for diagnosis, prognosis, and individualized treatment plans (Sanju Koirala et al., 2024).

To better represent the variety inherent in ADHD presentations, a fundamental movement away from standard category frameworks and toward dimensional diagnostic models is expected. At the same time, research is being done to create predictive biomarkers that can predict the onset, course, and response to treatment of the disorder, allowing for earlier and more successful interventions. Future studies to focus on bigger, more varied cohorts and longitudinal designs to evaluate and reinforce findings. Present methodological constraints are limited sample sizes and the prevalence of group-level analyses . Integrative methods that combine neurophysiological and measurements, such as electroencephalography, are also becoming more popular because they provide a thorough understanding of the structural and functional abnormalities of the brain linked to ADHD (Loo et al., 2025).

8. Conclusion

ADHD is caused by abnormal brain circuits and is a complicated neurological disorder. Prefrontal cortex, basal ganglia, and cerebellum experience disruptions in development that affect core cognitive and behavioral functions like attention, impulse control, and executive functioning.

There is no one-size-fits-all approach to diagnosis and therapy since ADHD is so diverse. As research on developmental trajectories continues, especially with longitudinal designs, more distinct trends on the best times and methods for interventions are becoming apparent. Dopamine and norepinephrine regulate behavior and attention and further support the necessity of behavioral and pharmaceutical strategies to address specific neurochemical imbalances. Identifying reliable biomarkers and embracing precision phenotyping are promising steps toward more tailored and evidence-based interventions. Challenges remain in translating these findings into routine practice, especially given the brain's dynamic nature and the individual variability seen across populations.

Moving past stereotypes and embracing the complexities of ADHD is essential to improving our understanding of the condition. Incorporating neuroscience, genetics, psychology, and empirical data can help future studies and clinical initiatives better assist people with ADHD by providing not only symptom relief but also a road to success. The changing nature of neuroscientific research and the variation in study methodology limit this work, even if it gives a general summary of how ADHD affects brain development and function. It is challenging to extrapolate results from small sample sizes or without long-term follow-up to a variety of demographics. Most research focuses on children and adolescents, with fewer studies examining how ADHD manifests in adulthood. Cultural, socioeconomic, and gender-based differences in diagnosis and treatment are also often overlooked.

It is possible to give priority to extensive longitudinal studies that track brain development across time, from early childhood to adulthood. To find out how genetic predispositions and environmental factors like stress, food, and digital media interact to affect ADHD symptoms, more research is needed.

Works Cited

Cunnington, Ross. “Neuroplasticity: How the Brain Changes with Learning.” IBE — Science of Learning Portal, 2019, solportal.ibe-unesco.org/articles/neuroplasticity-how-the-brain-changes-with-learning/.

da Silva, Bruna Santos, et al. “An Overview on Neurobiology and Therapeutics of Attention-Deficit/Hyperactivity Disorder.” Discover Mental Health, vol. 3, no. 1, 5 Jan. 2023, link.springer.com/article/10.1007/s44192-022-00030-1, https://doi.org/10.1007/s44192-022-00030-1

Ferranti, Anthony S, et al. “Novel Pharmacological Targets for GABAergic Dysfunction in ADHD.” Neuropharmacology, vol. 249, 1 Mar. 2024, pp. 109897–109897, https://doi.org/10.1016/j.neuropharm.2024.109897

Gao, Zhiyao, et al. “Reduced Temporal and Spatial Stability of Neural Activity Patterns Predict Cognitive Control Deficits in Children with ADHD.” Nature Communications, vol. 16, no. 1, 8 Mar. 2025, https://doi.org/10.1038/s41467-025-57685-x

Hassan, Sadaf. “A Clinician’s Guide to Understanding ADHD in 2025 | Qbtech.” Qbtech, 4 June 2025, www.qbtech.com/blog/clinicians-guide-to-adhd-in-2025

Jones, Heather . “Do ADHD Symptoms Differ in Boys and Girls?” Verywell Health, 17 Jan. 2022, www.verywellhealth.com/do-adhd-symptoms-differ-in-boys-and-girls-5207995

Juliano-Villani, Gabrielle . “The ADHD vs. Non-ADHD Brain: What Are the Differences?” ChoosingTherapy.com, 2023, www.choosingtherapy.com/adhd-brain-vs-normal-brain/

Lange, Klaus W., et al. “The History of Attention Deficit Hyperactivity Disorder.” ADHD Attention Deficit and Hyperactivity Disorders, vol. 2, no. 4, 30 Nov. 2010, pp. 241–255, www.ncbi.nlm.nih.gov/pmc/articles/PMC3000907/, https://doi.org/10.1007/s12402-010-0045-8

Loo, Sandra K, et al. “Translating Decades of Neuroscience Research into Diagnostic and Treatment Biomarkers for ADHD.” Advances in Neurobiology, 1 Jan. 2024, pp. 579–616, https://doi.org/10.1007/978-3-031-69491-2_20

Mamiya, Ping, et al. “Precision Medicine Care in ADHD: The Case for Neural Excitation and Inhibition.” Brain Sciences, vol. 11, no. 1, 13 Jan. 2021, p. 91, https://doi.org/10.3390/brainsci11010091

Mayo Clinic. “Adult Attention-Deficit/Hyperactivity Disorder (ADHD) - Symptoms and Causes.” Mayo Clinic, 25 Jan. 2023, www.mayoclinic.org/diseases-conditions/adult-adhd/symptoms-causes/syc-20350878

Michelini, Giorgia, et al. “Treatment Biomarkers for ADHD: Taking Stock and Moving Forward.” Translational Psychiatry, vol. 12, no. 1, 12 Oct. 2022, https://doi.org/10.1038/s41398-022-02207-2

National Institute of Mental Health. “Attention-Deficit/Hyperactivity Disorder (ADHD).” Www.nimh.nih.gov, National Institute of Mental Health, Dec. 2024, www.nimh.nih.gov/health/topics/attention-deficit-hyperactivity-disorder-adhd

Paus, T. S. A. A. “Structural Maturation of Neural Pathways in Children and Adolescents: In Vivo Study.” Science, vol. 283, no. 5409, 19 Mar. 1999, pp. 1908–1911, science.sciencemag.org/content/283/5409/1908, https://doi.org/10.1126/science.283.5409.1908

Polanczyk, Guilherme, et al. “The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis.” American Journal of Psychiatry, vol. 164, no. 6, 1 June 2007, pp. 942–948, ajp.psychiatryonline.org/doi/full/10.1176/ajp.2007.164.6.942, https://doi.org/10.1176/ajp.2007.164.6.942

Rubia, Katya. “Cognitive Neuroscience of Attention Deficit Hyperactivity Disorder (ADHD) and Its Clinical Translation.” Frontiers in Human Neuroscience, vol. 12, no. 100, 29 Mar. 2018, www.ncbi.nlm.nih.gov/pmc/articles/PMC5884954/, https://doi.org/10.3389/fnhum.2018.00100

Sanju Koirala, et al. “Neurobiology of Attention-Deficit Hyperactivity Disorder: Historical Challenges and Emerging Frontiers.” Nature Reviews. Neuroscience, 24 Oct. 2024, www.nature.com/articles/s41583-024-00869-z, https://doi.org/10.1038/s41583-024-00869-z

Silva, Sandra. “ADHD and Dopamine: What’s the Connection?” Healthline, 10 Oct. 2024, www.healthline.com/health/adhd/adhd-dopamine

Stiles, Joan, and Terry L. Jernigan. “The Basics of Brain Development.” Neuropsychology Review, vol. 20, no. 4, 3 Nov. 2010, pp. 327–348, link.springer.com/article/10.1007/s11065-010-9148-4, https://doi.org/10.1007/s11065-010-9148-4

UFO Themes. “Attention Deficit Hyperactivity Disorder: An Overview.” Neupsy Key, 13 June 2021, neupsykey.com/attention-deficit-hyperactivity-disorder-an-overview

Present Over Future: Prioritizing Moral Obligations to the Living in Policy-Making
By Hanna Yu

Introduction

While all human life holds inherent value regardless of when it exists, the moral obligations we owe to living persons differ significantly from what we owe future persons. The numerous differences between present and future people result in a difference in terms of treatment. We have a moral obligation to create systems and policies to address the identifiable needs of living persons that we do not owe future persons; our obligations to future persons are more abstract and should be largely rooted in safeguarding conditions for a viable future given the inherent uncertainty about their needs. Consequently, policy-making should prioritize the time-sensitive issues and interests of those alive today over the hypothetical problems future persons face.

Should We Give Equal Weight?

Considering the premise that all human life is valuable irrespective of when they live, it can be argued that lives that have yet to begin should not be assigned lesser moral weight than those alive today. Therefore, policy-makers should treat the needs of the two populations as equally important and should create frameworks that simultaneously address the needs of all.

However, in equating moral weights across time, we are inherently diminishing the numerous differences between the two populations and overlooking theories constructed around those differences. For example, the Will Theory brings forth a key distinction between the two populations in terms of having rights.[1] This theory is agency-centered and specifies that part of having rights is possessing the ability to exercise said rights. The capacity to take advantage of a right can only be possible for those currently alive. This idea directly contrasts with the beliefs of Intergenerational Justice theorists, who claim that rights do not require current agency. These theorists believe that though future people do not yet have the capacity to exercise their rights, they will be able to someday and should still be able to hold valid justice claims now. The rights of future people therefore deserve equal moral consideration, and policy makers who adopt the intergenerational justice perspective must create policies that reflect this. Rather than comparing the lives of present persons to future people, policy makers should take a non-comparative path such as a "threshold concept," which sets a minimum standard for acceptable well-being based on absolute conditions, not in relation to others.[2] Under this view, the harm threshold defines the point at which a person's well-being can drop below an acceptable "moral minimum". Even if future persons do not currently exist, they retain the potential to fall below this threshold, which grants them legitimate moral consideration even in present-day policy making. Considering that all people have an equal claim to reach this basic standard, instead of tying value to a temporal context, individuals of different times should be evaluated independently.

Yet despite this hypothetical scenario in which policy makers equate the moral values of present and future people, there are numerous limitations in doing so. For one, it is difficult to

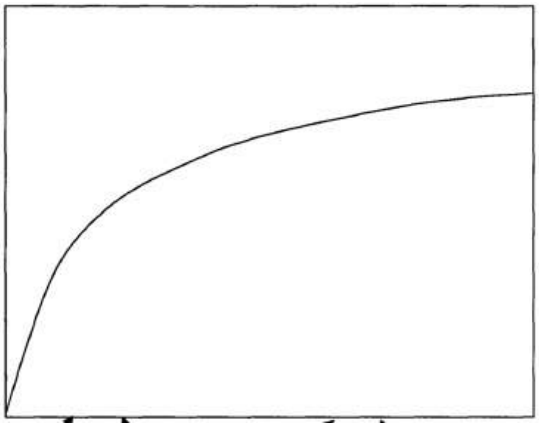
propose the moral values of the two populations as equal because future people, though projected to deal with harsher environmental conditions, also are predicted to have greater resources to combat their problems. Moreover, because policy-makers are inherently uninformed about the exact needs of future people, even in equating the moral values of present and future people they are still unable to create policies that address future people's problems.

The Prioritarian Argument

Derek Parfit's Priority View is a principle in moral philosophy that claims that helping people matters more when they are worse off, not with egalitarian motives of making everyone equal, but of giving extra moral importance to improving the lives of those who are worse off. The priority view, combined with projections indicating that future generations will most likely enjoy greater resources and thus possess the capacity to effectively mitigate any possible environmental or social challenges, provides a basis to support prioritizing the interests of those currently alive when making policy. Assuming policy is created with positive intents to hold citizens accountable, safeguard global resources, manage risks, and help lead nations toward a better future, policy should prioritize those currently alive.

When optimizing the well-being for a group of people, the idea of utilitarianism, maximizing the good and minimizing the bad, is often at the forefront of discussions. Utilitarianism, a moral theory aimed at maximizing well-being for the greatest number of people, theoretically seems to align with the intent behind policy making. However, while utilitarianism simply adds up utility, Prioritarianism cares not just about how much well-being is produced, but also about who receives it. This distinction is shown in the Utility Monster thought experiment, where one being derives significantly more utility from resources than everyone else.[3] While a utilitarian would be compelled to direct most resources to this monster, since it maximizes total utility even if it's at the expense of everyone else, a prioritarian would argue that helping the worse off has greater moral importance than helping the monster. Unlike total utilitarianism, prioritarianism gives greater moral weight to improving the wellbeing of those who are worse off, making this approach more attuned to issues of fairness of distributions. Prioritarianism can be graphed as an increasing and concave function, where the x-axis shows an individual's utility and the y-axis is the moral value policy-makers should ascribe to each person.[4]

Chart 1. Transforming Utility with a Strictly Increasing and Concave *g*-Function.



Because this function depicts the diminishing moral value of gains as well-being increases, well-being gains are worth more when they go to the worse off. As a policy-making framework, prioritarianism ensures that populations are not treated as interchangeable units of utility; rather, it emphasizes addressing the identifiable needs of those currently alive, while still considering the interests of future persons, though not weighing them equally.

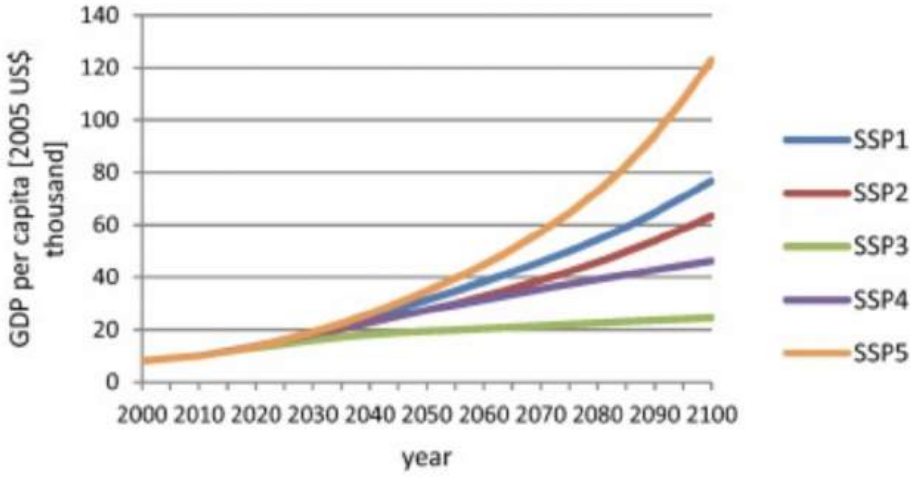
This distribution promotes a balanced approach to maximizing well-being, but it should also be applied because of the asymmetry of influence between present and future persons: the present generation has the unmatched ability to influence the behavior of future people by impacting their circumstances, or exercise power by jeopardizing the interests of future generations as well.[5] These unreciprocated dynamics place present people in a position of power over future people, and policy should prioritize the needs of them over the hypothetical needs of future persons as any effort directed toward present people also indirectly benefits the future. The investments policy makers put in the present is not shortsighted; it is a strategic way of safeguarding the future as well.

The "discount rate" of population ethics is a percentage used by economists and policymakers to reduce the value of future costs and benefits when making decisions, reflecting the assumption that events or consequences occurring one year from now are less important than those happening in the present.[6] Policy makers must consider the discount rate to ensure that the policies they craft, assuming their effects are felt across generations, reflect assumptions about future people's well-being. A discount rate is a way to mathematically adjust the value of future costs and benefits in today's terms, reflecting policy makers' time preference; in applying a discount rate, policy-makers effectively prioritize the present over the future, assigning less weight to the well-being of future persons.

The Uncertainty Argument

Because future outcomes are inherently uncertain, policy-makers cannot know which future-oriented policies would be effective. Policy-makers bear the responsibility to not only

prioritize the urgent needs of the current population, but also should either focus on initiatives that are almost guaranteed to have future benefits, such as growing the economy, or concentrate on the conditions necessary for a viable future. Projections for Gross Domestic Product (GDP) per capita are shown below; the different colors represent different Shared Socioeconomic Pathways (SSP), which are a set of scenarios used in climate research to describe possible future global developments in areas such as society, economics, technology, and the environment up to the year 2100.[7]



Though all of the pathways are hypothetical projections about the future, every scenario depicts a future that is significantly wealthier than the current world today; though policy makers can be fairly certain that global wealth will increase, we still cannot know the details of what the future will hold, making it difficult to justify creating policy aimed at speculative future benefits at the cost of the wellbeing of current generations.[8] It is important to note that predictions about the future often rest on fragile foundations of limited data, unobservable counterfactuals, and assumptions that are entirely unverifiable. For example, estimating the long-term effects of complex legislative changes or determining the impact of deterrence policies like the death penalty requires blending data with oftentimes untestable assumptions. As a result, policy analysis faces a trade-off between the strength and the credibility of its conclusions. While strong assumptions may yield definitive predictions, these conclusions are often less credible.[9]

Despite this tension, policy analysts are still often incentivized to express unwarranted certainty as both policy makers and the public typically prefer clear-cut answers over approaches that admit uncertainty. This tendency can distort decision making, leading policymakers to potentially overestimate the reliability of current knowledge. Given these limitations and risks, policy makers have the responsibilities to prioritize pressing current needs and focus on actions that would most likely produce long-term benefits, such as investing in economic growth or creating the foundational conditions necessary for a viable future.

Historically, efforts to establish a global institutional framework to protect future generations have largely fallen short in the field of international law. For example, despite

aspirations toward a unified global climate policy, policies remain fragmented, a reality acknowledged even by the Intergovernmental Panel on Climate Change (IPCC) under the heading "climate governance."[10] Though policymakers often utilize the rhetoric of "future generations," this language rarely translates into concrete, enforceable commitments.

Moreover, the idea that acting for "our" future generations will automatically benefit future generations everywhere is not only flawed but potentially harmful. There are many relevant disparities between countries: while wealthy countries may prioritize adaptation or slow mitigation because they expect to manage or even benefit from a moderate global warming, vulnerable nations such as small island states face existential threats that demand immediate global mitigation. The concept of "future generations" is therefore not a unified category as it masks the starkly unequal distribution of risks, sacrifices, and benefits across countries and populations; it is not only more realistic but possibly more just to prioritize the tangible needs of current, vulnerable populations rather than relying on vague appeals to future generations. Policy makers should reflect such in allocating fewer resources to uncertain future-oriented policies and instead prioritize the immediate, identifiable needs of the current population. While the wellbeing of future generations matters, the benefits of actions taken today for their sake are often indirect and unmeasurable, unlike the direct and measurable impact of policies on those alive today. Nearly all policy carries some degree of uncertainty, but the uncertainty increases as you look further into the future. The implication is not to ignore the future, but to act with caution to ensure sustainability, avoid irreversible harm, and safeguard future options; this calls for balancing present needs with future concerns without overcommitting resources to uncertain outcomes.

Conclusion

While future persons deserve moral consideration, the obligations we owe them differ drastically from those policy makers owe to living persons. The current population can benefit immediately from policy addressing their identifiable needs, while future persons can only be the indirect beneficiaries of present-day political action. Our moral obligations to living people are more concrete, whereas obligations to future people are uncertain; in policy-making, frameworks like prioritarianism and concepts such as the discount rate further justify why present needs deserve greater moral weight in decisions, not out of disregard for future lives, but because effective policy must act on what can be known, impacted, and measured. Prioritizing present populations ultimately safeguards future ones by ensuring the foundational conditions of stability, growth, and justice that future persons will inherit.

Works Cited

Franks, D. M., Davis, R., Bebbington, A. J., Ali, S. H., Kemp, D., & Scurrah, M. (2014). Conflict translates environmental and social risk into business costs. Global Environmental Change 30, 1–11. https://doi.org/10.1016/j.gloenvcha.2014.10.021

Gosseries, A. (2020). Intergenerational justice. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2020 Edition). https://plato.stanford.edu/entries/justice-intergenerational/

Nozick, R. (1974). Anarchy, State, and Utopia. Basic Books.

Purdy, J. (2015). The politics of nature: Climate change, environmental law, and democracy. Duke Law Faculty Scholarship. https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=5231&context=faculty_scholarship

Wilde, R. (2022). Climate change, international law and the 'intergenerational turn'. European Journal of International Law 33(4), 1061–92. https://doi.org/10.1093/ejil/chac059

Winkler, H., Dubash, N. K., & Rajamani, L. (2018). Facilitating an equitable energy transition in developing countries. Nature Climate Change 8(12), 1011–16. https://pubmed.ncbi.nlm.nih.gov/30478061/

Notes

1. Gosseries 2020.
2. Ibid.
3. Nozick 1974, p. 41.
4. Purdy 2015.
5. Gosseries 2020.
6. Wilde 2022.
7. Franks et al. 2014.
8. Gosseries 2020.
9. Winkler et al. 2018.
10. Wilde 2022.

Role of Social Determinants in Bladder Stimulation Therapy Access and Perceived Efficacy

By Karthik Ventrapragada

Abstract

Bladder control disorders affect hundreds of millions of people around the world and lead to countless physical, psychological, and social consequences. Although bladder stimulation therapies such as Sacral Neuromodulation (SNM), Vagus Nerve Stimulation (VNS), Percutaneous Tibial Nerve Stimulation (PTNS), and Pelvic Floor Electrical Stimulation (PFES) can offer promising support in managing bladder incontinence symptoms, their effectiveness relies on a combination of social determinants. This paper presents a survey of 112 individuals, most of whom are older adults who have experienced or currently experiencing bladder control issues. It explores how different variables such as age, access to healthcare, treatment costs, dependable transport, as well as social support affect the perceived efficacy of various bladder stimulation therapies. Overall, the outcomes show that most patients experience life-changing improvements in managing their symptoms and quality of life after treatment. However, external barriers such as high costs, insufficient insurance coverage, and countless others remain issues to achieve the best possible care. These insights highlight the need for healthcare policies that focus on helping patients with access to bladder stimulation to improve their outcomes and quality of life.

Keywords

Bladder stimulation therapy; Neuromodulation; Social determinants of Health; Healthcare accessibility; Quality of life; Incontinence; Sacral neuromodulation; Patient outcomes; Financial barriers; Transportation challenges

Introduction

1. Overview of Bladder Stimulation Therapy

Bladder control disorders are significant worldwide health issues, impacting millions of individuals worldwide and resulting in high medical expenses and poor quality of life (Irwin et al.; Abrams et al.). Symptoms range from urinary urgency, frequency, incontinence, and retention, which can impair a person's physical, mental, and social wellbeing (Irwin et al.). Common causes include neurological diseases such as multiple sclerosis and Parkinson's disease, spinal injuries, and age-related loss of bladder function (Kabay et al.). For most patients, not only is there physical discomfort, but also anxiety, embarrassment, and potentially social isolation.

Traditional methods of treating bladder disorders involve changing behavioral habits, taking medication, and exercising pelvic muscles (Booth et al.; Sayner et al.). Although these approaches may provide some relief, they may not work for everyone, particularly for those with strong and persistent symptoms. To address this, medical professionals have developed bladder stimulation techniques that utilize electrical impulses to regulate bladder function more effectively (Coolen et al.; Zhu et al.). These techniques target the nerves and muscles involved in urinary control and offer a potential solution for patients who have not found relief through

conventional therapy.

Bladder stimulation therapies enhance the communication between the bladder, spinal cord, and brain. They modulate nerve activity through electrical impulses to prevent irregular bladder contractions, improve muscle strength and coordination, and restore control over the bladder (Booth et al.). Although bladder stimulation therapies rely on their technology, much of their success depends on non-medical factors, such as accessibility to healthcare, financial costs, and patient perceptions of comfort and effectiveness (Sayner et al.; Ammirati et al.). Therefore, understanding both the technology, as well as the social factors influencing these therapies is important in order to improve patient care.

2. Types of Bladder Stimulation Therapies

This paper focuses on multiple bladder stimulation therapies. Each therapy has distinct applications and mechanisms to treat different cases of bladder control issues.

Figure 1. Common Stimulations

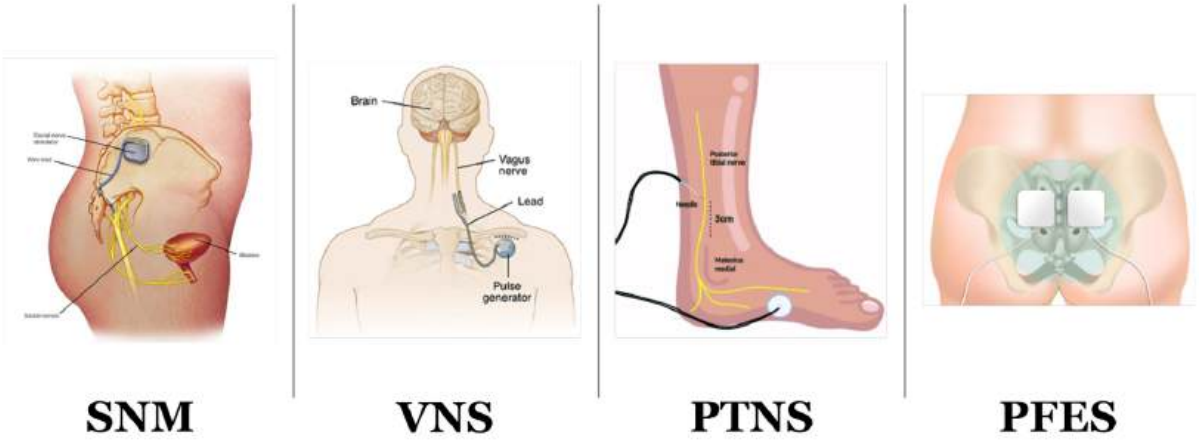


Figure 1 illustrates four common types of bladder stimulation therapies: Sacral Neuromodulation (SNM) ("Mayo Clinic"), Vagus Nerve Stimulation (VNS) ("Saint Luke's"), Percutaneous Tibial Nerve Stimulation (PTNS) (Khasanah et al.), and Pelvic Floor Electrical Stimulation (PFES) (Allon). Each method targets specific nerves or muscle groups to help regulate bladder function.

Sacral Neuromodulation (SNM)

Sacral Neuromodulation therapy involves surgically implanting a small device near the sacral nerves in the lower back (Coolen et al.; Laviana et al.). These nerves play a crucial role in regulating overall bladder function. During the procedure, a lead wire is placed near the sacral nerve and an electrical pulse generator is implanted under the skin. This generator sends small electrical impulses to the sacral nerves, helping to restore normal bladder control with electrical nerve signals (Coolen et al.)

SNM is typically used for patients with overactive bladder, urinary retention, or other

bladder control disorders that do not respond to traditional treatment (Coolen et al.; Ammirati et al.). Clinical trials on SNM have shown significant improvements in bladder control and quality of life for patients undergoing this therapy (Ammirati et al.; Laviana et al.).

Transcutaneous Electrical Nerve Stimulation (TENS)

Transcutaneous Electrical Nerve Stimulation is a non-invasive therapy that uses electrodes to deliver electrical impulses through the skin (Booth et al.; Sayner et al.). These impulses stimulate the nerves responsible for bladder control, reducing symptoms of urgency and incontinence. Unlike SNM, TENS does not require surgery and can be administered at home or in clinics.

Patients using TENS report symptom improvement after regular sessions (Booth et al.; Sayner et al.). However, the effectiveness of this therapy can vary based on factors such as electrode placement, stimulation frequency, and patient adherence to the treatment schedule.

Percutaneous Tibial Nerve Stimulation (PTNS)

Percutaneous Tibial Nerve Stimulation involves inserting a thin needle near the ankle to stimulate the tibial nerve, which connects indirectly to the sacral nerves (Kabay et al.). By targeting this nerve, PTNS can influence bladder control without the need for invasive surgery.

The procedure is usually performed in outpatient settings with sessions taking around 30 minutes. Patients often require multiple sessions over several weeks to achieve significant results. PTNS has been effective in managing symptoms of overactive bladder and incontinence in clinical studies (Burton et al.; Ahadi et al.; Sonmez et al.).

Pelvic Floor Electrical Stimulation (PFES)

Pelvic Floor Electrical Stimulation uses a probe inserted into the vagina or rectum to deliver electrical impulses to the pelvic floor muscles (Booth et al.). These muscles play a critical role in supporting the bladder and maintaining urinary control. Strengthening the pelvic floor can reduce symptoms of incontinence and improve bladder function.

PFES is often used along with pelvic floor exercises, such as Kegel exercises, to increase muscle strength and coordination. Patients typically have several sessions under the supervision of a healthcare provider.

Vagus Nerve Stimulation (VNS)

Vagus Nerve Stimulation is primarily used to treat epilepsy and depression but has also shown potential for improving bladder control by modulating the autonomic nervous system (Jezernik et al.). The vagus nerve influences various bodily functions, including bladder activity. While research on VNS for bladder control is still in its early stages, preliminary findings suggest that it may help regulate bladder contractions and reduce urgency.

Other Types of Electrical Stimulation Therapies (EST)

There are various other types of Electrical Stimulation Therapy which consist of various

techniques that use electrical currents to stimulate nerves and muscles involved in bladder control. EST can be administered through different devices and methods, including surface electrodes and implanted stimulators (Van Balken et al.). It aims to restore normal bladder function by improving nerve communication and muscle responsiveness.

3. Mechanisms of Bladder Stimulation Therapies

Bladder stimulation therapies operate on the principle of neuromodulation, or the manipulation of neural activity by utilizing electrical impulses (Coolen et al.; Zhu et al.). The sacral nerves, which have the responsibility for carrying messages between the bladder, spinal cord, and brain, are the primary nerves addressed with these therapies. In conditions such as overactive bladder or urinary retention, these nerve pathways can become interrupted, resulting in inappropriate bladder contractions or the inability to completely empty the bladder.

By delivering electrical impulses, these stimulation therapies help restore normal communication in the nervous system. The impulses have the ability to inhibit hyperactive nerve signals or stimulate hypoactive ones, depending on the patient's condition. Research shows that neuromodulation can induce long-term changes in the activity of nerves, providing long-term symptom relief even after treatment sessions are stopped (Zhu et al.; Ammirati et al.).

In addition to nerve-based therapies, treatments like Pelvic Floor Electrical Stimulation (PFES) focus on strengthening the coordination and power of pelvic muscles. The pelvic floor plays a crucial role in supporting the bladder and maintaining continence (Coolen et al.). The electrical stimulation of these muscles boosts their function, reducing leakage and improving bladder control.

The success of bladder stimulation treatments may vary from one patient to another. Factors such as nerve health, underlying medical conditions, and compliance with treatment regimens may influence the outcomes. Additionally, research also calls for personalized treatment plans, as different patients may respond better to certain types of stimulation.

4. Social and Environmental Factors Affecting Bladder Therapy

Access to healthcare services, including bladder stimulation therapies, is often shaped by a complex mix of social and environmental factors. These factors can create barriers that prevent patients from receiving effective treatment (Irwin et al.; Sayner et al.; Ammirati et al.). Understanding these barriers is essential to improving healthcare delivery and patient outcomes.

Healthcare Accessibility

Patients in rural or underserved areas may face significant challenges in accessing specialized care (Sayner et al.). Long travel distances to medical facilities can discourage patients to attend regular follow-up visits, particularly for therapies that require multiple sessions, such as PTNS and PFES. The availability of healthcare providers trained in bladder stimulation techniques also varies geographically, contributing to differences in care.

Transportation and Reliability

Reliable transportation is critical for attending medical appointments, especially for patients with mobility issues or those living far from healthcare facilities. Patients without access to personal vehicles may need to rely on public transport, rideshare services, or assistance from family members. Unreliable transportation can result in missed appointments, delaying or disrupting treatment plans.

Financial Costs and Insurance Coverage

The cost of bladder stimulation therapies can be a significant barrier for many patients (Sayner et al.; Ahadi et al.). Surgical procedures like Sacral Neuromodulation may involve high upfront costs, including device implantation and post-operative care. Even non-invasive therapies such as TENS and PTNS may require ongoing expenses for equipment and multiple treatment sessions. Insurance coverage plays a critical role in determining whether patients can afford these therapies. Inadequate coverage can lead to financial strain, forcing patients to abandon treatment.

Patient Comfort and Provider Communication

Comfort levels with discussing urinary symptoms and undergoing bladder stimulation procedures can influence willingness to seek care (Abrams et al.; Booth et al.). Cultural aspects, past healthcare experiences, and the perceived embarrassment surrounding bladder disorders can affect how openly patients communicate with providers. Healthcare professionals need to have a supportive and nonjudgmental environment to encourage patients to feel comfortable and engage in potential treatment (Abrams et al.).

5. Previous Studies on Effectiveness and Patient Outcomes

Numerous studies have demonstrated the benefits of bladder stimulation therapies in managing urinary symptoms (Booth et al.; Zhu et al.; Burton et al.). Patients often report reduced urinary frequency and urgency, fewer episodes of incontinence, and improved bladder emptying (Kabay et al.; Ahadi et al.; Sonmez et al.). These improvements can enhance overall quality of life by enabling greater participation in social activities within the community and reducing the burden associated with bladder control disorders.

Research also emphasizes the importance of consistent follow-up care and patient education (Sayner et al.; Ahadi et al.). Patients who understand their treatment options and follow prescribed protocols are more likely to achieve positive results. Also, social support from family members and healthcare providers can reinforce adherence to therapy, in turn promoting long-term success (Zhu et al.).

Despite these findings, some patients experience limited benefits or even adverse effects from bladder stimulation therapies (Coolen et al.; Ammirati et al.). Complications may include pain, discomfort, or infection at the placement/surgical site (Abrams et al.). Understanding these risks and working to reduce them is essential in optimizing patient care.

The purpose of this study is to investigate how social factors influence accessibility and

perceived effectiveness of various bladder stimulation therapies. This research attempts to close a gap in current literature by focusing on patient experiences instead of clinical outcomes alone. It explores how barriers such as healthcare access, transportation, finances, and health care provider communication affect both access to and perceived efficacy of these simulations. The results aim to inform healthcare practices to reduce the number of disparities in the field of electrical stimulations of the bladder, and improve quality of care for individuals with bladder control disorders.

Methodology

Survey Design

This study utilized a survey designed to assess social factors affecting the accessibility and perceived effectiveness of bladder stimulation therapies. This study utilized Google Forms for the structure and distribution of the survey, due to its ease for distribution and anonymous data collection. The survey included a combination of multiple-choice questions and Likert scale ratings. This study focused on participants' demographics, health conditions, experiences with bladder stimulation procedures, and perceived barriers to accessing and benefiting from these treatments.

The survey consisted of multiple parts, including demographics, health history, treatment experiences, accessibility and affordability, perceived effectiveness, side effects, and community participation, each one focused on one specific aspect of participants' experiences. The demographic section collected information on age, gender, race, education level, work situation, residence, and household income. The health-related section inquired about participants' neurological diagnoses, bladder control issues, and history of bladder stimulation treatment. Additional sections also asked about the accessibility, affordability, and perceived effectiveness of various therapies, as well as the side effects, and the troubles encountered.

Participants

A total of 112 participants voluntarily completed the survey. Respondents were recruited from diverse sources, including local hospitals, healthcare centers, and social media platforms such as Twitter/X, Reddit, and Facebook. This approach ensured a more comprehensive understanding of the social factors influencing both access to and outcomes of these treatments.

Participation was entirely voluntary, and respondents were not required to disclose their names or any personally identifiable information. The criteria specified that participants had to be at least 18 years old and capable of providing informed consent. There were no restrictions based on gender, race, or socioeconomic status, ensuring the sample reflected a diverse population.

Data Collection

Participants could access and complete the survey using any device able to connect to the internet, including smartphones, tablets, and computers. The approximate time for the completion of this survey was 15-20 minutes, and no incentives were provided for participation.

The data collection period lasted almost 12 weeks, allowing time to reach a broader audience.

To maintain the integrity of the data received, duplicate responses were removed based on timestamps. Furthermore, incomplete surveys were excluded from the final analysis, ensuring that only fully completed responses were included.

Ethical Considerations

Ethical guidelines were followed throughout the study. Participants gave informed consent before starting the survey and confirmed that they understood its purpose, that participation was voluntary, and that their responses would remain confidential. They were informed that they could withdraw at any time.

No personal identification was collected such as names, phone numbers, email addresses, etc. and all responses were stored securely using encrypted cloud storage. Access to the data was restricted to the primary researcher and advisor involved in the study.

Data Analysis

Once responses were collected, the data was exported to a spreadsheet format for analysis. Descriptive statistics such as frequencies, percentages, means, and standard deviations were used to summarize demographic information and survey responses. The goal was to identify general trends related to accessibility, perceived effectiveness, and patient-reported barriers to bladder stimulation therapies.

Findings were interpreted in the context of existing literature to ensure that conclusions were supported by both the results of this survey and prior research. This descriptive approach provided a broad understanding of the social factors that may influence access to and perceived effectiveness of bladder stimulation treatments.

Results

Participant Demographics

The survey included 112 participants, with an age range from 29 to 83 years and a mean age of 62.17±11.09 years. The gender distribution showed that 53.6% were female, 45.5% were male, and 0.9% identified as non-binary or other.

In terms of race, the majority identified as White with 42.9%, followed by Black or African American with 23.2% and Asian with 18.8%. Regarding education, 25.0% held a bachelor's degree, while 15.2% had completed high school. Employment status showed that 46.4% were retired, 16.1% were self-employed, and 11.6% were employed full-time. Table 1 summarizes the demographic characteristics of all participants.

Table 1. Patient Demographics

Variable	**Characteristics**	**Percent (Count)**
Age	20-40	4.5% (5)
	40-60	32.1% (36)

	60-80	59.8% (67)
	>80	3.6% (4)
Sex	Male	45.5% (51)
	Female	53.6% (60)
	Non-Binary	0.9% (1)
Race	White	42.9% (48)
	African-American/Black	23.2% (26)
	Asian	18.8% (21)
	Hispanic	13.4% (15)
	Native Hawaiian/Other Pacific Islander	0.9% (1)
	American Indian/Alaska Native	0.9% (1)
Education	Less than High School	0.9% (1)
	High school (including GED)	15.2% (17)
	Some college (no degree)	14.3% (16)
	Associate degree (2-year)	13.4% (15)
	Bachelor's degree (4-year)	25.0% (28)
	Technical certification	11.6% (13)
	Master's degree	13.4% (15)
	Doctoral degree	2.7% (3)
	Professional degree (JD, MD)	0.9% (1)
	Prefer not to say	2.7% (3)
Employment	Retired	46.4% (52)
	Self-employed	16.1% (18)
	Part-time employment	16.1% (18)

	Full-time employment	11.6% (13)
	Home-maker	7.1% (8)
	Unemployed	2.7% (3)
Marital Status	Married	46.4% (52)
	Widowed	20.5% (23)
	Separated	3.6% (4)
	Divorced	16.1% (18)
	Single/Unmarried	13.4% (15)
Income	Less than $25,000	0.9% (1)
	$25,000 - $50,000	30.4% (34)
	$50,000 - $75,000	25.9% (29)
	$75,000 - $100,000	23.2% (26)
	$100,000 - $125,000	9.8% (11)
	$125,000 - $150,000	6.2% (7)
	$150,000 - $200,000	1.8% (2)
	More than $200,000	1.8% (2)
Community type	Rural Area	29.5% (33)
	Suburban Area	58.0% (65)
	Urban Area	12.5% (14)
Household size	1 person	40.2% (45)
	2 people	30.4% (34)
	3 people	12.5% (14)
	4 people	13.4% (15)
	More than 4 people	3.6% (4)

Neurological Diagnoses and Stimulation Procedures

Participants reported a variety of neurological conditions, with stroke being the most

common at 26.8%, followed closely by spinal cord injury and multiple sclerosis both recorded at 25.0%. Conditions related to Parkinson's disease, brain or spinal tumors, and Alzheimer's disease were also frequently reported. Table 2 below provides a full breakdown of the neurological diagnoses reported by participants.

Table 2. Patient Neurological Diagnoses

(Participants could select multiple diagnoses. As a result, the total percentage exceeds 100%.)

Neurological Diagnosis	Percent (Count)
Stroke	26.8% (30)
Spinal Cord Injury (non-tumor)	25.0% (28)
Multiple Sclerosis	25.0% (28)
Parkinson's Disease	18.8% (21)
Brain or Spinal Cord Tumor	15.2% (17)
Alzheimer's Disease	15.2% (17)
Encephalitis	12.5% (14)
Cerebral Palsy	8.9% (10)
Traumatic Brain Injury (TBI)	6.3% (7)
Other	3.6% (4)

Sacral Neuromodulation (SNM) was the most commonly reported bladder stimulation therapy, used by 31.3% of participants. Other frequently reported therapies included Transcutaneous Electrical Nerve Stimulation (TENS) at 26.8%, Pelvic Floor Electrical Stimulation (PFES) at 24.1% and Percutaneous Tibial Nerve Stimulation (PTNS) at 17.9%. These therapies are among the most established methods for treating bladder dysfunction and were commonly selected across a diverse range of participants.

Some participants reported receiving multiple forms of stimulation therapy. The most frequent combination was SNM with other stimulation(s), with the combination occurring in 20 cases (17.9% of participants). Another common combination involved Other Electrical Stimulation Therapies (EST), appearing in 18 cases (16.1%). PFES combinations were also common as they were found in 14 responses (12.5%). PTNS-related combinations followed at 11 (9.8% of participants), and Vagus Nerve Stimulation (VNS) combinations were found in 8 participants (7.1%). These patterns suggest that many patients are either undergoing sequential treatments or combining therapies to improve outcomes.

Table 3 below provides a breakdown of all stimulation procedures reported by participants. Note: Table 3 does not show the different combinations, but rather the number of participants who reported each stimulation.

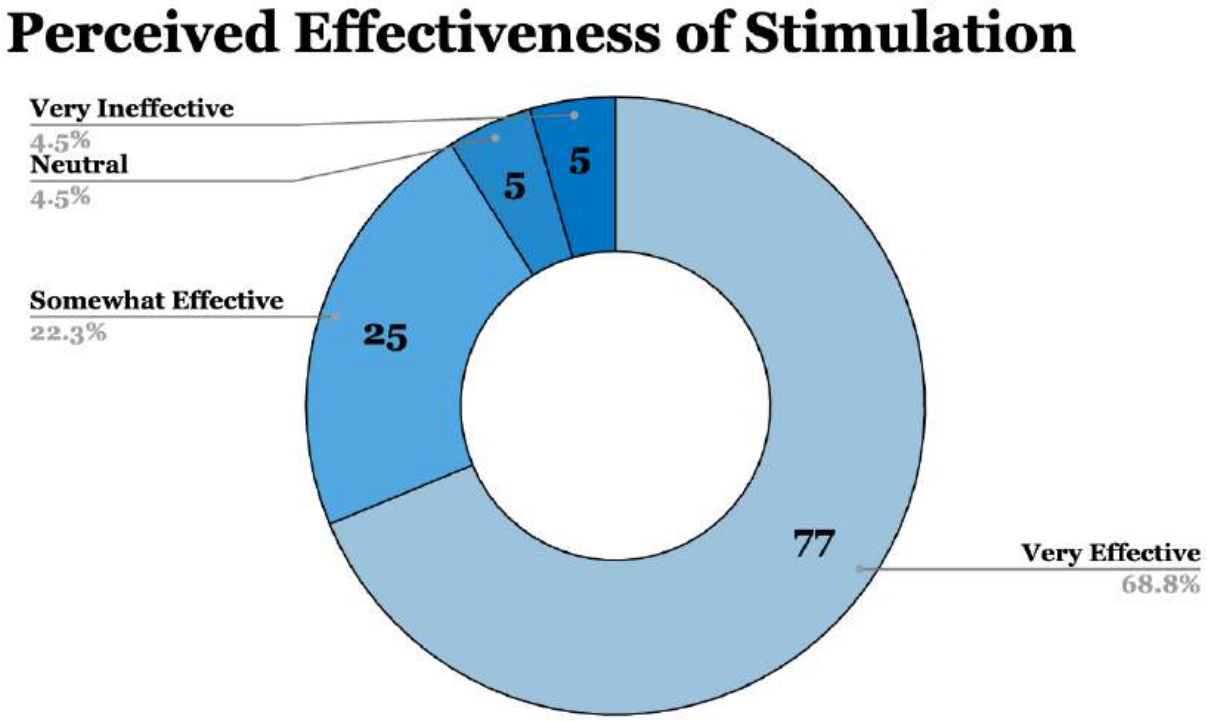
Table 3. Stimulation Procedure
(Participants could select multiple stimulations. As a result, the total percentage exceeds 100%.)

Stimulation Procedures	**Percent (Count)**
Sacral Neuromodulation (SNM)	31.3% (35)
Transcutaneous Electrical Nerve Stimulation (TENS)	26.8% (30)
Percutaneous Tibial Nerve Stimulation (PTNS)	17.9% (20)
Pelvic Floor Electrical Stimulation (PFES)	24.1% (27)
Other Electrical Stimulation Therapies (EST)	26.8% (30)
Vagus Nerve Stimulation (VNS)	14.3% (16)

Treatment Experience and Perceived Effectiveness

Participants were asked, “How effective do you find bladder stimulation in managing your bladder issues?” In response, 68.8% found bladder stimulation very effective, while 22.3% found it somewhat effective. Only 4.5% were neutral, and another 4.5% found it very ineffective. Figure 2 below highlights the distribution of responses regarding the perceived effectiveness of stimulation therapies.

Figure 2. Perceived Effectiveness of Stimulation



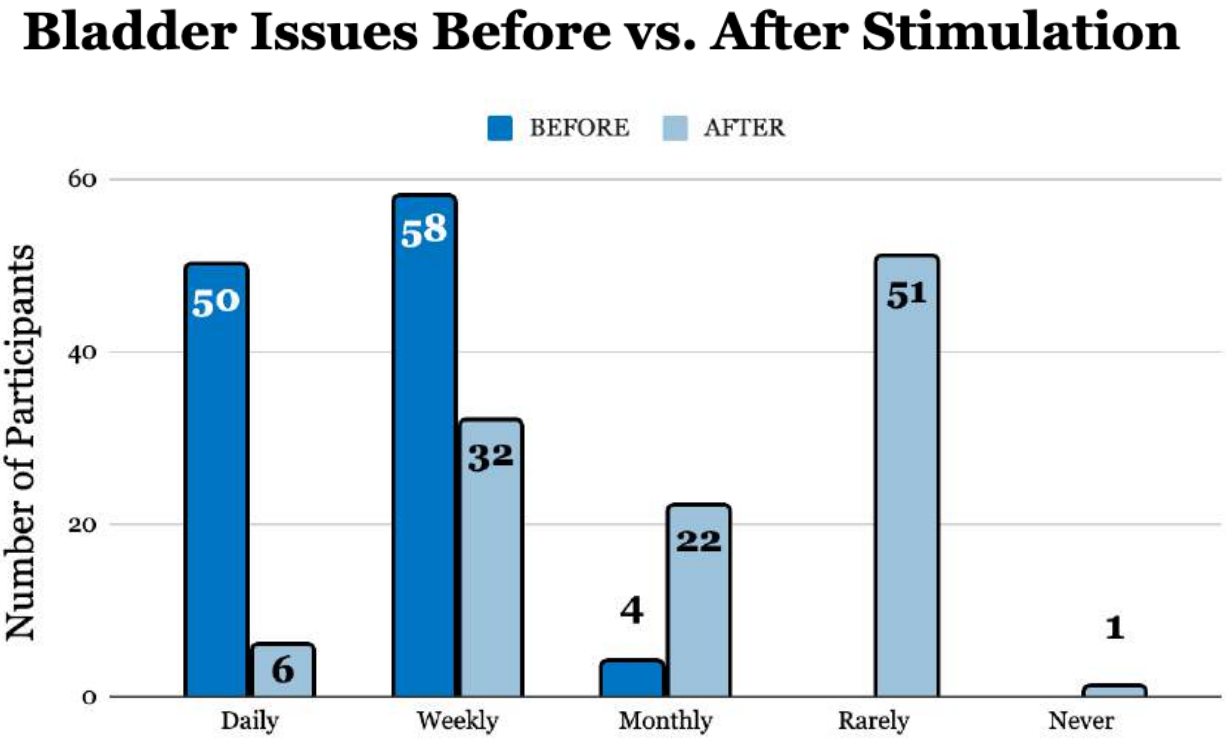
(Participants were given the options of Very Effective, Somewhat Effective, Neutral, Somewhat Ineffective, and Very Ineffective. The option Somewhat Ineffective was not chosen by any participants)

Participants were also asked, "Were there any complications or side effects with your stimulation procedure?" Most participants reported no side effects at 67.9% (76 participants). Among those who did, the most common complication was pain or discomfort at 30.4% (34 participants), followed by technology-related issues at 6.2% (7 participants) and persistent bladder symptoms also at 6.2% (7 participants). Note that participants could select multiple complications. As a result, the total percentage exceeds 100%**.**

Quality of Life Before and After Stimulation

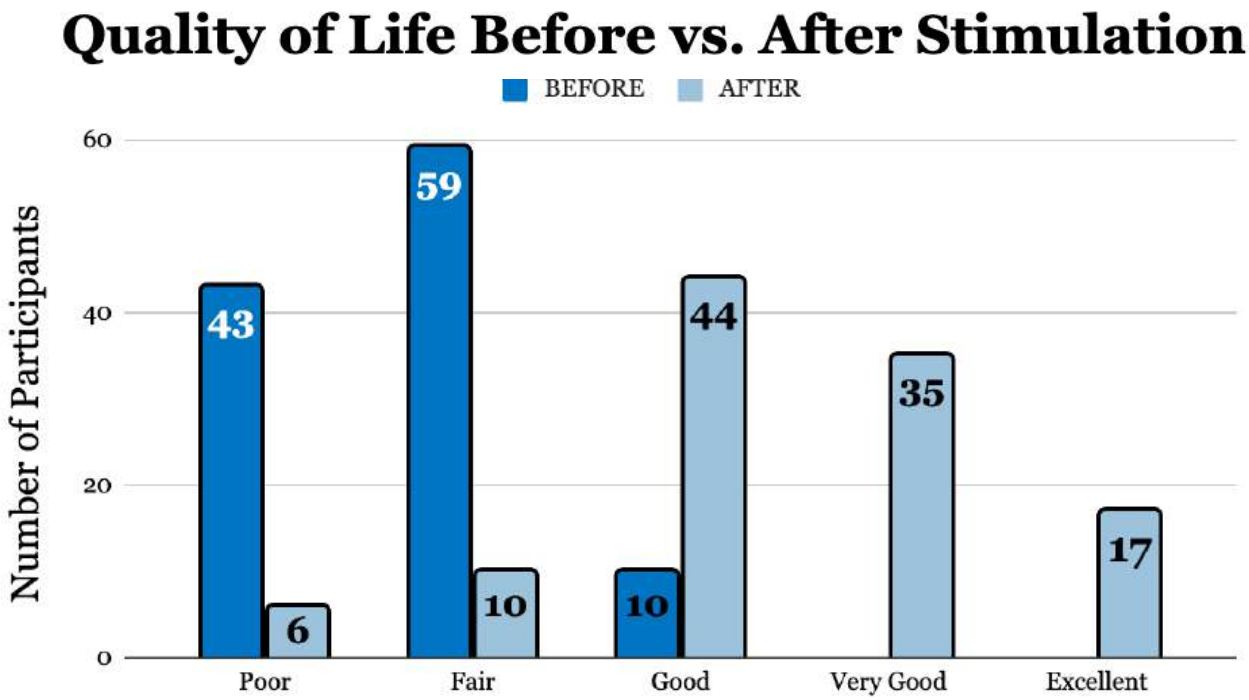
Participants were asked, "How often did you experience bladder control issues BEFORE your procedure?" and "How often did you experience bladder control issues AFTER your procedure?" Before stimulation, 44.6% experienced daily bladder control issues, and 51.8% experienced bladder control issues weekly. After treatment, 45.5% reported rarely experiencing bladder control issues, and only 5.4% continued to face bladder control issues daily. These changes are represented in Figure 3, which compares the frequency of bladder issues before and after stimulation.

Figure 3. Bladder Issues Before vs. After Stimulation



Participants were also asked, "How would you describe your quality of life BEFORE doing bladder stimulation?" and "How would you describe your quality of life AFTER doing bladder stimulation?" 38.4% rated their quality of life as poor before treatment. After treatment, only 5.4% rated it as poor, while 85.7% rated it as good, very good, or excellent, compared to just 8.9% who rated it as good before stimulation. Figure 4 shows how participants rated their quality of life before and after stimulation, with most reporting improvement after treatment.

Figure 4. Quality of Life Before vs. After Stimulation

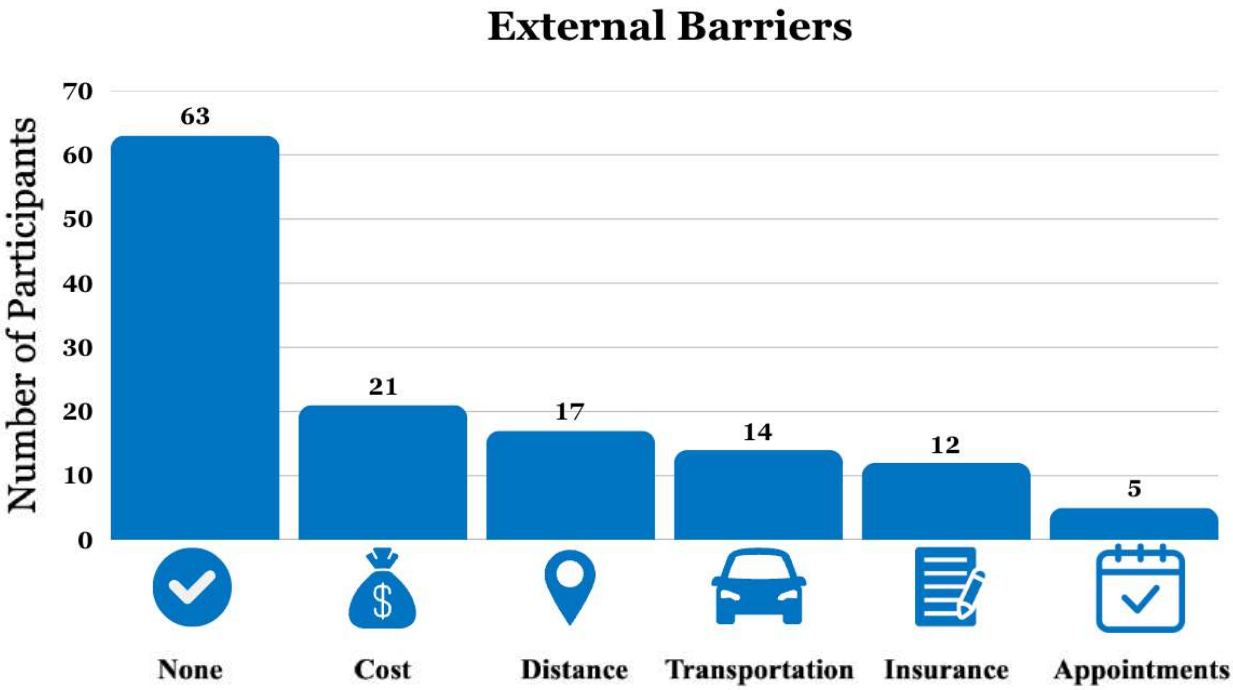


Accessibility and Challenges

Participants were asked, "How accessible was it for you to get bladder stimulation procedure?" In response, 55.4% of participants found the procedure very accessible, 24.1% reported neutral accessibility, and 2.7% found it somewhat inaccessible.

They were also asked, "What challenges have you faced in accessing bladder stimulation treatment?" From this, 18.8% of participants cited cost as a significant barrier. Insurance coverage issues were reported by 10.7%, and 15.2% indicated that distance to providers was a challenge. Figure 5 below shows the main external barriers participants faced when trying to access treatment.

Figure 5. External Barriers
(Participants could select multiple diagnoses. As a result, the total percentage exceeds 100%.)



Participants were given the options of None, Cost of Treatment, Distance to Healthcare, Transportation, Insurance Coverage, and Availability of Appointments.

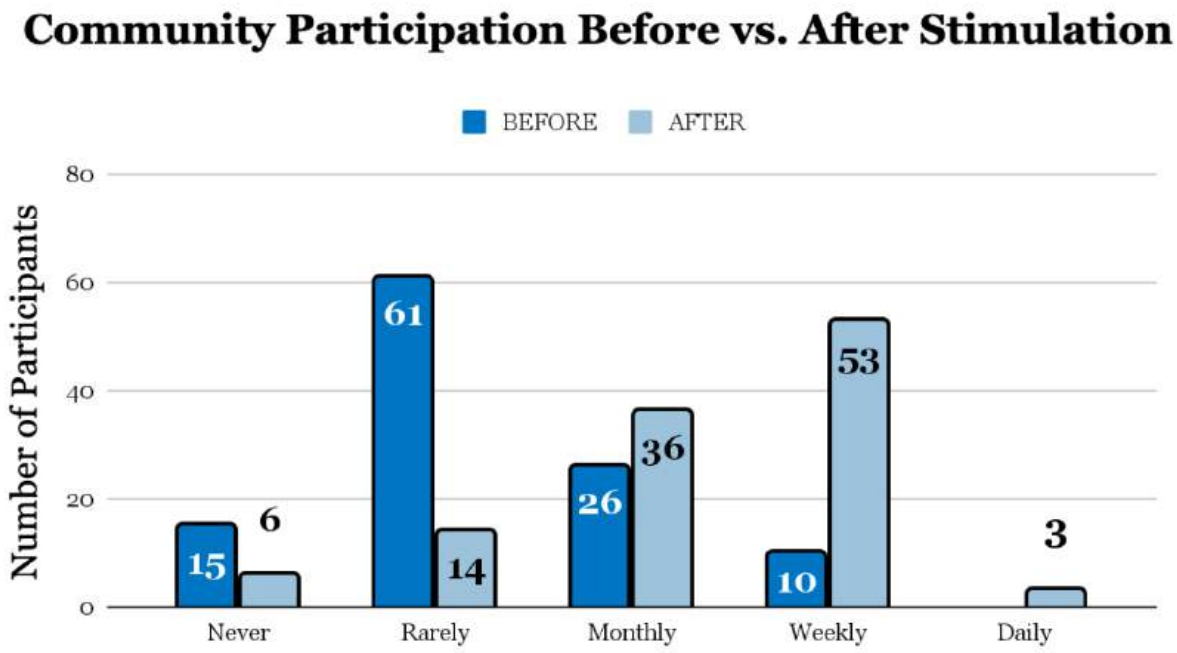
Transportation and Healthcare Accessibility

Participants were asked, "How far is your Urologist from your home?" In response, 12.5% lived less than 1 mile away, 9.8% lived 1–5 miles away, 38.4% lived 6–10 miles away, 11.6% lived 11–20 miles away, and 27.7% lived more than 20 miles from their urologist. Participants were also asked, "How do you typically get to your Urologist?" Most reported using a personal vehicle at 44.6%, followed by friends or family at 20.5%, public transportation at 18.8%, walking or biking at 13.4%, and taxi or rideshare services such as Uber or Lyft at 2.7%.

Social and Community Participation

Before treatment, 54.5% of participants rarely engaged in community activities. After treatment, 47.3% reported weekly participation. Participants were asked, "How comfortable do you feel in public places before and after receiving bladder stimulation treatment?" In response, 60.7% reported feeling much more comfortable in public, 28.6% felt slightly more comfortable, and 10.7% reported no change. Figure 6 shows how participants' levels of community participation changed before and after receiving bladder stimulation treatment.

Figure 6. Community Participation Before vs. After Stimulation



Discussion

The demographic findings show that bladder stimulation therapy is most commonly used by older adults, with the average participant age being 62.17. This finding is in agreement with existing knowledge that bladder dysfunction increases with age due to factors such as pelvic floor muscle weakening, reduced nerve sensitivity, and conditions like Parkinson's disease and diabetes (Irwin et al.). The high percentage of retired individuals of 46.4% also supports research showing that urinary symptoms often interfere with daily life in later years which prompts older adults to seek more permanent treatments like neuromodulation (Kabay et al.).

The high prevalence of neurological conditions, such as spinal cord injuries at 25%, multiple sclerosis also at 25%, and stroke at 26.8%, among the participants underscores the link between neurological impairment and bladder dysfunction. This relationship is well researched and documented, and studies show that damage to the brain-bladder axis leads to disrupted bladder control (Van Balken et al.). The popularity of Sacral Neuromodulation (SNM) in this group follows evidence supporting its use in more difficult cases (Jezernik et al.; Kabay et al.). On the other hand, the underuse of non-invasive options like Percutaneous Tibial Nerve Stimulation (PTNS) and Transcutaneous Tibial Nerve Stimulation (TTNS) might be due to limited provider familiarity or availability, as found in various systematic review studies (Booth et al.; Burton et al.).

The overwhelmingly positive results, with 91.1% of participants reporting the therapy as very or somewhat effective, support the current clinical consensus and confirm that bladder stimulation is effective, especially for treating overactive bladder and incontinence. (Irwin et al.; Laviana et al.). However, the 30.4% of people who reported pain or discomfort shows that

patients may not be getting enough information before treatment and that there needs to be better education about side effects and what to expect (Coolen et al.).

The reliability of the technology appears to be satisfactory, with only 6.3% reporting malfunctions. This also confirms findings that highlight the high success of bladder stimulation devices in the long term (Zhu et al.). However, other literature expresses that specific devices and various patient characteristics can also influence outcomes, highlighting the need for more specialized research in future studies (Van Balken et al.).

Bladder stimulation was also strongly linked with improved quality of life. The drop in self-reported poor quality of life, from 38.4% to 5.4%, supports findings that bladder dysfunction deeply affects emotional well-being, confidence, and daily functioning (Ahadi et al.). The increase in weekly community participation from 8.9% to 47.3% post-treatment is in agreement with findings showing that bladder therapies help patients reintegrate socially and reduce anxiety (Ammirati et al.).

Despite these benefits, accessibility remains to be a major challenge. Nearly 30% of participants reported financial obstacles, highlighting concerns about inconsistent insurance coverage and high costs, reported at 10.7% and 18.8% respectively.

Transportation was another key factor in treatment access. With 21.4% of patients depending on public transport or rideshare options, current studies show the importance that inconsistent transport can lead to missed sessions, especially for time-intensive treatments like PTNS which require weekly appointments over the course of several weeks (Coolen et al.).

The improvement in public comfort, 60.7% of participants felt much more confident in public after their treatment and further underscores the psychological burden of bladder dysfunction. This finding reinforces that the impact of these therapies extends beyond physical symptoms, improving emotional and social health (Ahadi et al.; Burton et al.).

Limitations

This study can offer many important insights to healthcare providers looking to improve patient outcomes and barriers to treatment, however there are a few limitations that should be considered. Firstly, the data came from a self-reported survey so there is always a chance that some participants gave inaccurate or biased responses. Another limitation is that the study did not ask questions to differentiate between the effects of different simulation therapies, which may have influenced participants' responses and overall therapy effectiveness.The survey was also only available online, which may have excluded some people without internet access or tech familiarity. These limitations do not take away from the findings but they do show that future research could go deeper and include a more diverse range of participants and data.

Conclusion

This study shows that though bladder stimulation therapies work extremely well for most patients, their effectiveness is strongly dependent on social determinants like access to care, transportation, cost, and ability to communicate with providers. The findings emphasize that

numerous individuals experience significant benefits in symptoms and quality of life following treatment but that not all are able to benefit equally because of external obstacles. These results indicate that healthcare systems must do better to eliminate those obstacles and make bladder stimulation therapies more accessible. Future studies should target specific therapies and have more diverse populations of patients to enhance outcomes for everyone.

Acknowledgement Statement

I would like to thank my mentor Amanda Rounds, PhD and Indigo Research for helping me and guiding me through this project. Their support made a big difference in my research. I also confirm that the ideas, writing, and graphics in this paper are completely my own.

Works Cited

Abrams, P., et al. “Overactive Bladder Significantly Affects Quality of Life.” The American Journal of Managed Care, vol. 6, no. 11 Suppl, July 2000, pp. S580-590.

Ahadi, Tannaz, et al. “Efficacy of Percutaneous vs Transcutaneous Posterior Tibial Nerve Stimulation in Overactive Bladder Syndrome: A Randomized Clinical Trial.” Basic and Clinical Neuroscience, vol. 15, no. 4, 2024, pp. 499–508. PubMed, https://doi.org/10.32598/bcn.2023.4896.1.

Allon, Eleanor. “The Role of Neuromuscular Electrical Stimulation in the Rehabilitation of the Pelvic Floor Muscles.” British Journal of Nursing, 8 Aug. 2019, https://www.britishjournalofnursing.com/content/focus/the-role-of-neuromuscular-electrical-stimulation-in-the-rehabilitation-of-the-pelvic-floor-muscles/.

Ammirati, Enrico, et al. “Neuromodulation in Urology, State of the Art.” Urologia, vol. 86, no. 4, Nov. 2019, pp. 177–82. PubMed, https://doi.org/10.1177/0391560319866075.

Booth, Joanne, et al. “The Effectiveness of Transcutaneous Tibial Nerve Stimulation (TTNS) for Adults with Overactive Bladder Syndrome: A Systematic Review.” Neurourology and Urodynamics, vol. 37, no. 2, Feb. 2018, pp. 528–41. DOI.org (Crossref), https://doi.org/10.1002/nau.23351.

Burton, C., et al. “Effectiveness of Percutaneous Posterior Tibial Nerve Stimulation for Overactive Bladder: A Systematic Review and Meta-Analysis.” Neurourology and Urodynamics, vol. 31, no. 8, Nov. 2012, pp. 1206–16. PubMed, https://doi.org/10.1002/nau.22251.

Coolen, Rl, et al. “Electrical Stimulation in the Treatment of Bladder Dysfunction: Technology Update.” Medical Devices: Evidence and Research, vol. Volume 12, Sept. 2019, pp. 337–45. DOI.org (Crossref), https://doi.org/10.2147/MDER.S179898.

Irwin, Debra E., et al. “Population-Based Survey of Urinary Incontinence, Overactive Bladder, and Other Lower Urinary Tract Symptoms in Five Countries: Results of the EPIC Study.” European Urology, vol. 50, no. 6, Dec. 2006, pp. 1306–15. DOI.org (Crossref), https://doi.org/10.1016/j.eururo.2006.09.019.

Jezernik, Saso, et al. “Electrical Stimulation for the Treatment of Bladder Dysfunction: Current Status and Future Possibilities.” Neurological Research, vol. 24, no. 5, July 2002, pp. 413–30. PubMed, https://doi.org/10.1179/016164102101200294.

Kabay, Sahin, et al. “The Clinical and Urodynamic Results of Percutaneous Posterior Tibial Nerve Stimulation on Neurogenic Detrusor Overactivity in Patients With Parkinson’s Disease.” Urology, vol. 87, Jan. 2016, pp. 76–81. DOI.org (Crossref), https://doi.org/10.1016/j.urology.2015.09.026.

Khasanah, Nurida, et al. “Physical Agent-Based Treatments for Overactive Bladder: A Review.” ResearchGate. www.researchgate.net, https://doi.org/10.3390/jcm11175150. Accessed 17 July 2025.

Laviana, Aaron, et al. “Sacral Neuromodulation for Refractory Overactive Bladder, Interstitial Cystitis, and Painful Bladder Syndrome.” Neurosurgery Clinics of North America, vol.

25, no. 1, Jan. 2014, pp. 33–46. PubMed, https://doi.org/10.1016/j.nec.2013.08.001.
"Mayo Clinic." Sacral Neuromodulation in the Male Patient, https://www.mayoclinic.org/medical-professionals/urology/news/sacral-neuromodulation-in-the-male-patient/mac-20549082. Accessed 17 July 2025.
"Saint Luke's." Understanding Vagus Nerve Stimulator (VNS) Placement, https://www.saintlukeskc.org/health-library/understanding-vagus-nerve-stimulator-vns-placement. Accessed 17 July 2025.
Sayner, Alesha M., et al. "Transcutaneous Tibial Nerve Stimulation in the Management of Overactive Bladder: A Scoping Review." Neuromodulation, vol. 25, no. 8, Dec. 2022, pp. 1086–96, https://doi.org/10.1016/j.neurom.2022.04.034.
Sonmez, Rafet, et al. "Efficacy of Percutaneous and Transcutaneous Tibial Nerve Stimulation in Women with Idiopathic Overactive Bladder: A Prospective Randomised Controlled Trial." Annals of Physical and Rehabilitation Medicine, vol. 65, no. 1, Jan. 2022, p. 101486. DOI.org (Crossref), https://doi.org/10.1016/j.rehab.2021.101486.
Van Balken, Michael R., et al. "THE USE OF ELECTRICAL DEVICES FOR THE TREATMENT OF BLADDER DYSFUNCTION: A REVIEW OF METHODS." Journal of Urology, vol. 172, no. 3, Sept. 2004, pp. 846–51. DOI.org (Crossref), https://doi.org/10.1097/01.ju.0000134418.21959.98.
Zhu, De Ting, et al. "Therapeutic Effects of Electrical Stimulation on Overactive Bladder: A Meta-Analysis." SpringerPlus, vol. 5, no. 1, Dec. 2016, p. 2032. DOI.org (Crossref), https://doi.org/10.1186/s40064-016-3737-5.

The Impact of Virtual Reality in Physical and Neurological Rehabilitation By Shay Bhatia

Abstract

Virtual reality (VR) technologies are increasingly being integrated into neurorehabilitation, offering novel ways to engage patients and accelerate recovery. This literature review examines the role of VR in enhancing physical, cognitive, and psychological outcomes across diverse rehabilitation contexts. Drawing from recent studies, the review is organized into three core domains: physical rehabilitation, where VR improves motor function and adherence; virtual embodiment and psychological well-being, where immersive technologies simulate bodily presence to restore disrupted body awareness, enhance the sense of agency, and reduce psychological distress in patients with neurological injuries; and cognitive recovery, demonstrating improvements in attention, memory, and executive function through immersive interventions. Collectively, these findings underscore VR's capacity to reshape traditional rehabilitation through multisensory feedback, gamification, and neuroplastic engagement, summarized in Fig 1. The review concludes by exploring future research directions, including expanding VR applications beyond clinical settings and addressing barriers to accessibility and standardization.

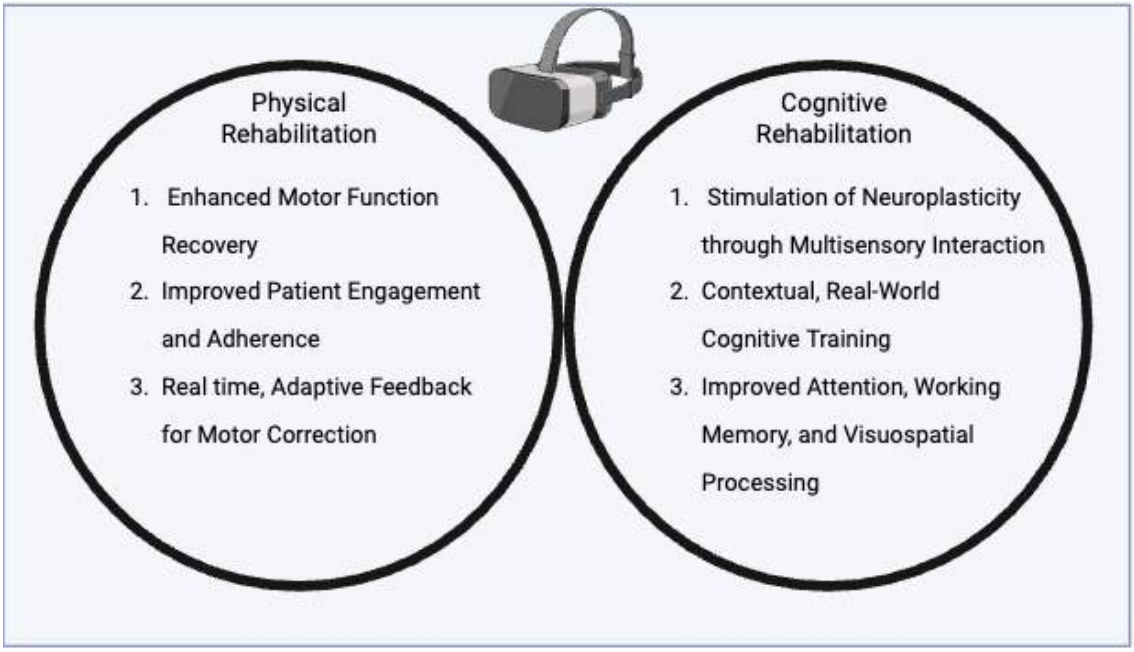


Fig 1: Visual representation of the benefits posed by VR-based rehabilitation in physical and cognitive rehabilitation (Created in BioRender)

Introduction

Virtual Reality (VR) is an immersive technology that uses three-dimensional environments, typically experienced through head-mounted displays (HMDs) and motion tracking devices, allowing users to navigate and interact with simulated worlds. In the context of

healthcare and rehabilitation, VR has gained increasing attention as an innovative tool for physical recovery and neurorehabilitation. This increase in attention is driven by its ability to engage patients in highly immersive and customizable therapeutic exercises that are not possible in traditional rehabilitation methods (Borrego et al.; Świdrak et al.).

Over the past decade, VR technology has exhibited promising outcomes across physical, cognitive, and psychological areas of rehabilitation. For physical rehabilitation, VR has been shown to enhance functional recovery, improve motor learning, and increase patient adherence to the rehabilitation process through the gamification of tasks made possible by VR and immersive feedback mechanisms. These advancements are especially significant in the recovery of motor function following injuries or immobilization, where VR enables patients to visualize and rehearse specific movements, thereby stimulating neuroplasticity in the absence of physical execution (Combalia et al.; Borrego et al.; Saby et al.)

A notable feature of VR in rehabilitation is the concept of virtual embodiment (VE), in which users experience a sense of ownership, agency, and control over a virtual body or limb. Beyond physical outcomes, VR has also demonstrated significant cognitive benefits in patients recovering from neurological injuries such as a stroke and traumatic brain injury. Immersive VR systems have enabled stroke survivors to practice daily life activities in realistic, safe, and gamified environments. This approach to recovery is associated with improvements in memory, attention, executive functioning skills, as well as reduced anxiety and pain, and shorter hospital stay lengths (Chatterjee et al.). These cognitive improvements highlight VR's potential to address post-stroke cognitive impairment, which is an area traditionally underserved in rehabilitation programs. Despite these encouraging findings, challenges and limitations remain in the widespread adoption of VR within clinical practice. Barriers such as high costs, technical limitations, and the need for robust longitudinal studies persist today (Akbar et al.). Addressing these challenges is vital to fully realizing VR's potential in enhancing patient outcomes and integrating it into standard care.

In this review, we will explore how virtual reality technologies can enhance physical, cognitive, and psychological outcomes in neurorehabilitation and physical recovery. Section 1 will focus on the application of VR in physical rehabilitation, highlighting multiple functional and motor benefits. Section 2 will examine virtual embodiment, its role in rehabilitation through VR, and its impact on body perception and pain modulation. Section 3 will discuss the cognitive benefits of VR, with a focus on its applications in neurorehabilitation. Finally, we will discuss the implications of these findings, the remaining challenges and limitations in VR for rehabilitation, and propose future directions for research and clinical integration.

Section 1: Applying Virtual Reality to Physical Rehabilitation

Virtual reality (VR) has emerged as a groundbreaking tool in physical rehabilitation, offering patients immersive, interactive environments that transform how therapeutic exercises are delivered and experienced. By combining real-time motion tracking, first-person perspectives, and engaging gamified elements, VR helps overcome common barriers in

rehabilitation, such as lack of motivation, fear of pain, and reduced adherence, while improving functional outcomes (Schuermans et al.; DiCesare et al.). These systems are not merely passive distractions; they actively engage the brain and body, supporting recovery for a wide range of conditions, from post-surgical immobilization to athletic injuries.

The following subsections examine how VR enhances physical rehabilitation by addressing immobilization challenges, improving sports recovery and performance, elevating functional outcomes, and fostering greater patient engagement and adherence through gamification.

1.1. Overcoming Immobilization and Physical Limitations:

One of the most formidable challenges in rehabilitation occurs when patients are immobilized due to injury, surgery, or pain. VR provides a unique solution by allowing patients to rehearse movements and engage motor circuits even when actual movement is impossible. In a study by Trujillo et al., patients recovering from chronic lower back pain participated in seven VR sessions lasting 20–45 minutes. They controlled a virtual avatar performing graded exposure tasks such as virtual dish stacking and target shooting: movements that had long been avoided due to fear of pain. The intervention led to significant reductions in pain intensity and improvements in pain-related psychological symptoms (Trujillo et al.).

M-Gomez et al. reinforced these findings in their study of patients with complex regional pain syndrome (CRPS). Using a head-mounted display, participants embodied a virtual avatar of their affected limb and performed movements that were otherwise too painful or impossible in real life. The intervention not only reduced subjective pain levels but also activated motor-related brain areas. The authors emphasized that this visual-motor coupling supports neuroplasticity by engaging neural pathways required for action planning, making VR a powerful adjunct to conventional rehabilitation. (Matamala-Gomez, Diaz Gonzalez, et al.)

1.2. Supporting Sports Recovery and Enhancing Performance:

Beyond immobilization, VR has gained momentum in sports rehabilitation, where it serves both recovery and performance enhancement roles. In a systematic review, Cariati et al. examined six randomized controlled trials regarding virtual reality and sports performance. The populations included football players, basketball players, and healthy adults. They found that five out of the six randomized controlled trials reported positive effects on balance, sprint, and jump performance, neurocognitive function, reaction time, and technical sports skills when comparing the VR-enhanced training methods to traditional training methods. (Cariati et al.)

1.3. Improving Functional Outcomes Through Immersive Therapy:

As patients progress through rehabilitation, achieving functional independence remains a primary goal. VR interventions have shown measurable success in improving physical outcomes across diverse patient populations. Hoffman et al. reviewed VR-based therapy for acute procedural pain, hand rehabilitation, and VR embodiment therapy for phantom limb pain. They

found that seeing and controlling a virtual version of the missing limb can reduce phantom limb pain by counteracting maladaptive brain plasticity. (Hoffman et al.; Kikkert et al.; Matamala-Gomez, Donegan, et al.; Ramachandran and Rogers-Ramachandran). They concluded that VR is a promising nonopioid adjunct for acute pain and hand rehabilitation, and that VR has the potential to extend into chronic pain management, such as phantom limb pain.

A systematic review of 15 studies supported these findings, revealing that VR interventions resulted in greater improvements in balance scores, limb function, muscle activity and strength, as well as physical fitness and body composition. (Qian et al.)

1.4. Increasing Patient Engagement and Adherence:

Patient engagement is a critical determinant of rehabilitation success, yet many patients struggle to maintain motivation throughout lengthy therapy regimens. VR addresses this challenge by creating environments that are both enjoyable and immersive, making therapy feel less like a chore. In a study of 220 older adults, researchers reported high adherence rates among participants in VR-based training. Combalia et al. also recognized the increased adherence to therapy through VR-based training, where they noted VR's potential to eliminate the barrier of fear-avoidant behaviors, as well as VR's potential to tailor training to each patient, increasing the adherence to treatment. (Kwan et al.; Combalia et al.)

1.5. Harnessing Gamification for Rehabilitation Success:

Gamification within VR, embedding levels, challenges, and rewards, serves as a catalyst for patient motivation and adherence. Kwan et al. demonstrated that patients performing gamified VR tasks showed faster rehabilitation progression and higher motivation compared to those in non-gamified therapy. Schuermans et al. echoed these findings in orthopedic patients, where gamified VR interventions led to reduced pain perception and quicker achievement of therapy milestones. (Kwan et al.; Schuermans et al.)

1.6. Conclusion

By merging immersive environments, real-time feedback, and gamification, VR has redefined the possibilities of physical rehabilitation. It allows patients to rehearse movements despite immobilization, accelerates functional recovery, and increases adherence through engaging, interactive experiences. As research continues to demonstrate VR's efficacy across varied patient populations, it is becoming an indispensable component of modern rehabilitation programs, one that empowers patients to take an active role in their own recovery (Gsangaya et al.; Álvarez De La Campa Crespo et al.)

Section 2: Virtual Embodiment

Virtual embodiment refers to the phenomenon where individuals experience a sense of ownership and agency over a computer-generated body or body part within an immersive virtual environment. This illusion is constructed through the integration of multisensory cues, which

together convince the brain that the virtual body is, in some sense, its own (Kilteni et al.; Świdrak et al.). Importantly, embodiment in VR is not a single construct but is comprised of three key components: body ownership (a sense of identification with the virtual body, as if it belongs to oneself), agency (the sense of control over movements of the virtual body), and self-location (the perception of being spatially situated where the virtual body is located). This multisensory integration is compelling in immersive systems that offer first-person perspectives and real-time motion tracking, allowing users to see and move within a virtual world as if it were their own physical space (Matamala-Gomez, Donegan, et al.). In the context of rehabilitation, virtual embodiment has gained attention as a novel mechanism for enhancing both physical and neurological recovery. In neurorehabilitation, it serves as a tool to activate and reorganize motor and sensory circuits disrupted by injury or disease. For instance, patients with spinal cord injuries can engage with their virtual avatars to mentally rehearse movements, stimulating cortical regions involved in motor planning and execution even when physical movement is not possible (Mateo et al.). Similarly, in physical rehabilitation, embodiment enhances patient motivation and engagement by creating a compelling, interactive experience that encourages adherence to therapeutic exercises and reduces fear-avoidant behaviors. (Świdrak et al.)

Beyond motor function, virtual embodiment has also been shown to influence pain perception and emotional states. Studies have reported reductions in chronic pain (e.g., phantom limb pain) and improvements in psychological outcomes, such as confidence and agency, when patients interact with virtual bodies that are aligned with their movements (Combalia et al.; Matamala-Gomez, Diaz Gonzalez, et al.). We will explore in detail how virtual embodiment contributes to rehabilitation by addressing both physical implications, such as motor circuit activation and movement visualization, and mental implications, including restoration of body awareness, enhancement of agency, and modulation of pain.

2.1. Physical Implications

2.1.1. Activates Motor Circuits

Embodiment in virtual environments has been shown to stimulate motor areas of the brain, even in the absence of actual physical movement. This activation supports neuroplasticity, a crucial process for recovery in patients with motor deficits. Such stimulation primes the motor cortex, facilitating later physical rehabilitation by establishing stronger neural pathways. (Matamala-Gomez, Donegan, et al.; Mateo et al.)

2.1.2. Enables Movement Despite Physical Limitations

For patients unable to physically perform specific movements, virtual embodiment enables mental rehearsal and visualization of these actions. Studies demonstrate that imagining and observing movements through a virtual avatar can enhance motor planning and prepare the nervous system for actual movement. In their study, Matamala-Gomez et al. explored how virtual

embodiment affects patients with conservatively managed distal radius fractures. Participants were immersed in a virtual environment where they viewed and controlled a virtual avatar of their affected limb, allowing them to perform movements that were otherwise too painful or impossible in reality. The key findings were that the immersive VR group demonstrated significantly greater improvements in functional ability and range of motion (ROM) compared with the other two non-immersive VR groups. These improvements in functional ability and ROM were positively correlated with patients' self-reported feelings of ownership over the virtual arm displayed in the treatment. Notably, the study concluded that virtual embodiment is not merely a psychological distraction but has direct effects on sensorimotor processing and can act as a training mechanism for preparing the nervous system for actual physical movement. (Matamala-Gomez, Slater, et al.)

Similarly, Świdrak et al. extended these findings by demonstrating that virtual embodiment can also recalibrate disrupted body representations and engage sensorimotor processes in patients with fibromyalgia (FM), a population characterized by chronic pain and motor disturbances. In their 2023 study, Świdrak et al. investigated the effects of virtual embodiment on body perception, pain, and motor system engagement in women with FM. The researchers immersed 20 participants in a virtual environment, where they embodied either a fully visible virtual body or a progressively invisible one. During the visible body condition, participants performed simple movements (e.g., stretching arms, looking at their hands) while observing the avatar from a first-person perspective, creating a strong illusion of ownership and agency. The study found that participants who experienced embodiment of the visible virtual body reported enhanced motor engagement and a stronger sense of control over their movements. Those with greater disturbances in body perception (a common symptom in FM), showed a higher proneness to virtual embodiment, suggesting that their nervous systems were particularly receptive to recalibrating through multisensory feedback. Świdrak et al. emphasized that embodiment in VR engages sensorimotor processes by providing congruent visual and proprioceptive cues, even when physical movement is limited due to pain or fatigue. This mechanism supports motor planning and neural activation, which can be critical for preparing the nervous system to execute fundamental movements in later stages of rehabilitation. (Świdrak et al.)

2.2. Mental Implications of Virtual Embodiment

2.2.1. Restores Body Awareness and Spatial Cognition

Neurological injuries often disrupt patients' internal body maps and spatial orientation, resulting in impaired proprioception and difficulties navigating their environment. Virtual embodiment provides a unique method for recalibrating these disrupted representations through synchronized multisensory feedback. In Świdrak et al.'s study, female fibromyalgia patients underwent two VR sessions where they viewed a life-sized virtual body from a first-person perspective while performing synchronous hand movements. In one condition, the virtual body

remained fully visible; in the other, it progressively became transparent. Patients with higher baseline disturbances in body perception (FBAQ scores) reported stronger sensations of embodiment in the visible condition. These results demonstrate how embodiment recalibrates disrupted sensorimotor processing, offering a pathway to restore bodily awareness in patients who often see limited success with traditional therapies (Świdrak et al.). Borrego et al. further underscored VR's potential in this domain through a comparative study involving 32 post-stroke patients and 46 healthy controls. Participants engaged in a VR stepping task, where they were instructed to step on virtual targets using a gender-matched avatar from either a first-person perspective (head-mounted display) or a third-person perspective (screen-based). Results showed significantly higher body ownership and self-location in stroke survivors during first-person VR compared to third-person. This study highlights how immersive VR can reintegrate the sense of body in space even in patients with long-standing deficits. By restoring disrupted body maps through synchronized visual and proprioceptive feedback, VR interventions may not only facilitate neuroplastic changes critical for motor recovery but also enhance patient engagement and confidence in their rehabilitation process. These findings underscore VR's unique ability to address both the physical and psychological barriers to recovery that often persist in conventional therapies. (Borrego et al.)

2.2.2. Enhances Sense of Agency

The sense of agency, or the perception of controlling one's body and actions, is fundamental for emotional and psychological recovery after an injury. Combalia et al. highlighted how immersive VR environments reinforce this agency. Patients recovering from movement-related shoulder pain viewed a first-person avatar with synchronous upper body movement. They also observed a therapist avatar demonstrate shoulder movements. Patients triggered avatar movements by pressing a foot pedal, inducing agency without actual movement. Matamala-Gomez et al. expanded on this concept with a randomized controlled trial involving 54 patients who were immobilized after distal radius fractures. Participants were assigned to one of three groups: immersive VR (n = 20), non-immersive VR (n = 14), or conventional digit mobilization therapy (n = 20). The immersive VR group embodied a virtual arm and performed tasks like virtual grasping and object rotation for 15-minute sessions, five days a week, over two weeks. Results showed significantly greater improvements in functional ability and range of motion, which positively correlated with self-reported agency and ownership over the virtual limb. This study demonstrated that immersive VR significantly enhanced functional recovery and motor engagement compared to both non-immersive and traditional therapies due to the heightened sense of agency experienced by participants in the immersive VR group. (Combalia et al.; Matamala-Gomez, Slater, et al.)

2.2.3. Increases Cortical Excitability

Virtual embodiment also primes the nervous system for motor learning by increasing cortical excitability. In a study of 30 healthy adults, Matamala-Gomez et al. examined the effects

of different VR embodiment conditions on motor cortical activation. Participants observed hand movements while embodying either a full-body avatar, an upper-limb avatar, or a detached hand, with each condition lasting 10 minutes. The full-body and upper-limb conditions led to greater corticospinal activation, which was closely linked to stronger feelings of ownership and agency. This suggests that immersive, congruent embodiment conditions can effectively "awaken" motor pathways, preparing patients for more active rehabilitation. Trujillo et al. offered further support through a case series of two chronic lower back pain patients using Karuna Virtual Embodiment Training. Each completed seven VR sessions (~20–45 minutes) involving motor imagery and graded exposure tasks such as virtual dish stacking and target shooting. Both patients experienced significant pain relief and a reduction in fear of movement, gaining confidence and a greater sense of control through immersive VR. (Matamala-Gomez, Frisco, et al.; Trujillo et al.)

2.2.4. Reduces Pain and Emotional Distress

Beyond its cognitive and motor benefits, virtual embodiment also reduces pain and emotional distress by immersing patients in engaging environments that divert attention from pain and alleviate anxiety. Combalia et al. noted that VR's capacity for distraction complemented its motor benefits, providing a holistic approach to rehabilitation. Similarly, Matamala-Gomez et al. found that patients in the immersive VR group reported significantly lower disability and pain scores compared to those in the non-immersive and conventional groups. Chatterjee et al. demonstrated these emotional benefits in a randomized controlled trial of 40 stroke patients. This study showed that stroke patients using the VIRTUE VR system for tasks like virtual cooking, shopping, and household management experienced improvements in cognitive function, reduced anxiety, and higher satisfaction compared to a control group. They also had shorter hospital stays, suggesting that VR can support both emotional well-being and physical recovery. (Matamala-Gomez, Slater, et al.; Combalia et al.; Chatterjee et al.)

Section 3: Cognitive Benefits

3.1 Cognitive Recovery and Neuroplasticity

Virtual reality-based interventions have emerged as powerful tools for enhancing cognitive function across a range of domains: attention, memory, executive functioning, and visuospatial skills. These gains are particularly significant in populations recovering from stroke or traumatic brain injury, where neuroplasticity plays a critical role in rehabilitation. In one systematic review of over 1,500 chronic stroke patients, researchers found that VR significantly improved cognitive function, especially when delivered in short-term, intensive programs under eight weeks in duration (Gao et al.). These effects were even more pronounced when VR was integrated with conventional therapies, suggesting VR therapies provide the most benefit as a supplement to conventional therapies. Patients with severe cognitive impairments often face the steepest climb during recovery, yet immersive VR may offer them the most significant benefit. In a clinical trial using gamified simulations of daily activities, stroke patients receiving VR-based

therapy showed notable improvement in cognitive assessments, particularly in domains like attention and orientation. Remarkably, those with the most severe impairments made the largest gains, and they were also discharged from the hospital more quickly than their counterparts in a control group (Chatterjee et al.). Engaging with lifelike virtual tasks not only enhances cognitive metrics but also appears to reshape the brain itself. VR-based cognitive-motor training has been linked to increases in cortical gray matter and heightened beta wave activity: physiological signals of brain plasticity and engagement (Georgiev et al.). These biological markers are supported by functional improvements as well. Older adults who participated in immersive VR cognitive simulations that mirrored everyday tasks, such as shopping or cooking, showed significant boosts in global cognitive scores, executive function, and even physical frailty measures, all while experiencing very few side effects (Kwan et al.). The immersive nature of VR, with its ability to individualize difficulty levels and simulate real-life conditions, is what sets it apart. Studies have shown that this format activates mirror neurons and demands high levels of cognitive engagement, mechanisms known to drive plastic changes in the brain (Sokołowska). Rather than relying solely on traditional repetitive tasks, VR encourages users to solve problems and adapt in real time, making it a uniquely powerful modality in cognitive recovery.

3.2 Mental Training and Motor Imagery (MI)

When physical movement is restricted, whether due to injury, immobilization, or early-stage rehabilitation, mental rehearsal through Motor Imagery (MI) becomes a critical intervention. VR-enhanced MI deepens this effect by pairing visualization with sensory-rich environments that stimulate the brain's motor networks. Patients with cervical spinal cord injuries who engaged in imagined reach-to-grasp exercises in VR not only improved their coordination but also maintained these gains for months after the intervention ended. The key mechanism appeared to be the strengthened activation of motor regions responsible for planning and initiating movement (Mateo et al.). Even when delivered in non-immersive settings, MI has been shown to mitigate the adverse effects of immobilization. Participants who practiced imagining movement during a period of upper-limb suspension retained more function than those who received no intervention. However, the authors noted that immersive VR could further amplify these results through a combination of both motor imagery and action observation, something that is not possible outside of VR-based therapies (Combalia et al.). Together, these studies illustrate that mental training through VR can serve as a bridge, preserving and even enhancing cognitive-motor pathways during periods when physical movement is not yet possible.

3.3 Psychological Benefits and Real-World Outcomes

3.3.1. Reduced Anxiety, Depression, and Fatigue

Recovery from neurological trauma often carries a significant psychological burden. Feelings of depression, anxiety, and mental fatigue can interfere with rehabilitation progress and

overall quality of life. Fortunately, immersive VR offers relief on this front as well. Rather than simply distracting users, these systems provide deeply engaging, reward-driven environments that promote emotional well-being. In several studies, users participating in interactive VR exercise programs reported reduced tension and depressive symptoms, alongside improvements in mental clarity and energy levels (Qian et al.). The emotional benefits extend to more targeted clinical populations as well. In stroke rehabilitation, immersive VR was linked to significantly lower anxiety scores among patients with mild to moderate impairments. These effects were especially notable in patients who initially reported high levels of emotional distress (Chatterjee et al.). Motivational factors also appear to play a role: VR's gamified and interactive formats keep users engaged, which can reduce mental fatigue and boost emotional resilience throughout recovery. (Schuermans et al.)

3.3.2. Safe, Controlled, and Acceptable Rehab Environments

One of the most consistent findings across studies is that VR-based rehabilitation is not only practical but highly acceptable to patients and clinicians alike. In multiple clinical settings, immersive VR was reported as safe, with minimal side effects and high satisfaction rates. Stroke patients completing cognitive and motor tasks in calm, lifelike virtual environments rated their experience as highly enjoyable, with no adverse events reported (Fregna et al.). Similarly, in an elderly cohort undergoing cognitive training, fewer than 3% of sessions produced any symptoms of motion sickness or discomfort (Kwan et al.). These findings are critical in populations that may be hesitant to try new technology or have limited experience with digital tools. The strong tolerability and user acceptance suggest that VR can be successfully scaled to broader clinical use, provided barriers like cost and equipment accessibility are addressed (Akbar et al.)

3.3.3. Real-World Transfer and Long-Term Effects

Perhaps most importantly, the benefits of VR-based rehab appear to transfer beyond the clinic and into daily life. In one compelling case study, a military veteran suffering from persistent post-concussion symptoms showed dramatic recovery after completing a two-phase VR program. Designed to mimic both civilian and combat scenarios, the intervention restored his cognitive performance, balance, and visual motion tolerance, enabling him to return to full duty and resume his pre-injury activities (Rábago and Wilken). Similar real-world outcomes have been noted in larger trials as well. Stroke patients undergoing VR therapy demonstrated improved performance in activities of daily living (ADLs) (Gao et al.). These improvements were not just clinically significant; they translated to faster discharge times and reduced dependency on hospital care (Chatterjee et al.). In populations with mild cognitive impairment (MCI), immersive VR training enhanced visuospatial function, and virtual cognitive assessments provided more ecologically valid insights than traditional paper-based tests (Kang et al.; Liu et al.). These findings suggest that immersive VR doesn't just enhance recovery in artificial lab settings; it prepares users for the complexity and demands of everyday life.

Discussion

Virtual reality is rapidly evolving from a novel tool into a clinically impactful modality across rehabilitation. This review explored its use across three interconnected domains, physical rehabilitation, embodiment, and cognitive-psychological recovery, and found consistent evidence that immersive VR systems can enhance patient outcomes by leveraging neuroplasticity, engagement, and individualized feedback. More importantly, these domains do not function in isolation; the cognitive, emotional, and physical systems of the brain interact deeply during recovery, and VR is uniquely positioned to stimulate all three simultaneously.

In the context of physical rehabilitation, immersive VR systems have demonstrated improvements in motor performance, range of motion, gait, balance, and strength. These effects were consistently observed across patient populations: from stroke survivors to individuals recovering from orthopedic injuries (Corbetta et al.; Matamala-Gomez, Slater, et al.). Not only did patients regain physical function more quickly, but they also reported lower pain levels and higher satisfaction compared to those undergoing traditional therapy (Combalia et al.). The use of gamified and personalized VR programs played a key role in maintaining patient motivation, which in turn enhanced adherence and long-term engagement (Schuermans et al.). VR is thus not simply a delivery system; it is an active modifier of the rehabilitation experience.

The sense of embodiment, the feeling that the virtual body is one's own, emerged as a pivotal mechanism connecting physical outcomes with psychological transformation. Through virtual embodiment, patients not only engaged more actively in therapy but also reconstructed disrupted internal body maps. This was particularly evident in studies where first-person perspective and body continuity were manipulated to influence agency and motor cortical excitability (Borrego et al.; Matamala-Gomez, Frisco, et al.). In stroke survivors and patients with immobilization, embodiment helped restore the sense of ownership and control over one's body, enabling more effective motor recovery and improved spatial awareness. (Świdrak et al.; Mateo et al.)

Equally compelling were the cognitive and psychological outcomes. VR-based interventions consistently led to improvements in executive function, memory, and attention across stroke, spinal cord injury, and MCI populations (Gao et al.; Kwan et al.; Chatterjee et al.). These gains were not only statistically significant but in many cases functionally meaningful: patients returned to daily living tasks sooner, reported reduced cognitive fatigue, and even experienced shorter hospital stays. Underlying these improvements are mechanisms such as mirror neuron activation, dual-task processing, and mental imagery, all of which are effectively triggered by immersive and adaptive VR environments (Sokołowska; Clark et al.). The psychological benefits extended further. Several studies highlighted how immersive VR reduces anxiety, depression, and fear avoidance, especially in patients undergoing painful or challenging rehabilitation (Qian et al.; Chatterjee et al.). Importantly, patients described VR not merely as tolerable, but as enjoyable and empowering. High satisfaction scores and minimal side effects, reported even among older adults and those with cognitive impairments, support VR's feasibility in real-world clinical deployment (Fregna et al.; Kwan et al.)

Despite these strengths, several limitations remain. Small sample sizes, short durations, and heterogeneity in hardware and protocols limit many studies. As Akbar et al. notes, the field lacks rigorous longitudinal trials that assess the durability and cost-effectiveness of VR-based interventions. Moreover, high equipment costs and technical complexity may limit access in under-resourced clinics, raising concerns about health equity and scalability. (Akbar et al.)

Looking ahead, the use of virtual reality in rehabilitation and broader healthcare contexts should be guided by rigorous, large-scale studies that evaluate long-term outcomes across diverse patient populations. Standardized protocols and consistent outcome measures will be essential to scale VR effectively across medical, therapeutic, and wellness domains. Technological advances will continue to shape VR's role. Integrating artificial intelligence, wearable biosensors, or brain–computer interfaces could personalize interventions in real time, whether for physical therapy, cognitive training, or behavioral health. Home-based systems and mobile VR platforms also offer promising avenues for expanding access to care, education, and mental health services, especially in underserved or remote communities. As VR evolves, its applications will likely extend beyond clinical recovery into preventive care, public health education, and even workforce training, demanding a cross-disciplinary approach that combines medical, technological, and ethical perspectives.

In summary, this review underscores VR's potential not as a replacement for traditional rehabilitation but as an integrative platform that enhances engagement, accelerates recovery, and restores agency across multiple domains of function. Its ability to merge motor, cognitive, and psychological stimulation into a single, immersive experience marks a new chapter in rehabilitation science, one in which therapy becomes not only more effective, but more human-centered.

Works Cited

Akbar, Khalida, et al. "Reviewing the Horizon: The Future of Extended Reality and Artificial Intelligence in Neurorehabilitation for Brain Injury Recovery." Information, vol. 15, no. 8, Aug. 2024, p. 501. DOI.org (Crossref), https://doi.org/10.3390/info15080501.

Álvarez De La Campa Crespo, Mercè, et al. "Virtual Embodiment for Improving Range of Motion in Patients with Movement-Related Shoulder Pain: An Experimental Study." Journal of Orthopaedic Surgery and Research, vol. 18, no. 1, Sept. 2023, p. 729. DOI.org (Crossref), https://doi.org/10.1186/s13018-023-04158-w.

Borrego, Adrián, et al. "Embodiment and Presence in Virtual Reality After Stroke. A Comparative Study With Healthy Subjects." Frontiers in Neurology, vol. 10, Oct. 2019. Frontiers, https://doi.org/10.3389/fneur.2019.01061.

Cariati, Ida, et al. "Virtual Reality and Sports Performance: A Systematic Review of Randomized Controlled Trials Exploring Balance." Frontiers in Sports and Active Living, vol. 7, Apr. 2025. Frontiers, https://doi.org/10.3389/fspor.2025.1497161.

Chatterjee, Kausik, et al. "Immersive Virtual Reality for the Cognitive Rehabilitation of Stroke Survivors." IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, 2022, pp. 719–28. IEEE Xplore, https://doi.org/10.1109/TNSRE.2022.3158731.

Clark, Brian C., et al. "Adaptations in Human Neuromuscular Function Following Prolonged Unweighting: II. Neurological Properties and Motor Imagery Efficacy." Journal of Applied Physiology, vol. 101, no. 1, July 2006, pp. 264–72. DOI.org (Crossref), https://doi.org/10.1152/japplphysiol.01404.2005.

Combalia, A., et al. "Immersive Virtual Reality in Orthopaedics—a Narrative Review." International Orthopaedics, vol. 48, no. 1, Jan. 2024, pp. 21–30. DOI.org (Crossref), https://doi.org/10.1007/s00264-023-05911-w.

Corbetta, Davide, et al. "Rehabilitation That Incorporates Virtual Reality Is More Effective than Standard Rehabilitation for Improving Walking Speed, Balance and Mobility after Stroke: A Systematic Review." Journal of Physiotherapy, vol. 61, no. 3, July 2015, pp. 117–24. DOI.org (Crossref), https://doi.org/10.1016/j.jphys.2015.05.017.

DiCesare, Christopher A., et al. "High-Risk Lower-Extremity Biomechanics Evaluated in Simulated Soccer-Specific Virtual Environments." Journal of Sport Rehabilitation, vol. 29, no. 3, Mar. 2020, pp. 294–300. DOI.org (Crossref), https://doi.org/10.1123/jsr.2018-0237.

Fregna, Giulia, et al. "A Novel Immersive Virtual Reality Environment for the Motor Rehabilitation of Stroke Patients: A Feasibility Study." Frontiers in Robotics and AI, vol. 9, Aug. 2022. Frontiers, https://doi.org/10.3389/frobt.2022.906424.

Gao, Yong, et al. "Effects of Virtual Reality-Based Intervention on Cognition, Motor Function, Mood, and Activities of Daily Living in Patients With Chronic Stroke: A Systematic Review and Meta-Analysis of Randomized Controlled Trials." Frontiers in Aging Neuroscience, vol. 13, Dec. 2021. Frontiers, https://doi.org/10.3389/fnagi.2021.766525.

Georgiev, Danko D., et al. "Virtual Reality for Neurorehabilitation and Cognitive Enhancement."

Brain Sciences, vol. 11, no. 2, 2, Feb. 2021, p. 221. www.mdpi.com, https://doi.org/10.3390/brainsci11020221.

Gsangaya, Muhindra Rao, et al. "Comparison between the Effect of Immersive Virtual Reality Training versus Conventional Rehabilitation on Limb Loading and Functional Outcomes in Patients after Anterior Cruciate Ligament Reconstruction: A Prospective Randomized Controlled Trial." Asia-Pacific Journal of Sports Medicine, Arthroscopy, Rehabilitation and Technology, vol. 34, Oct. 2023, pp. 28–37. ScienceDirect, https://doi.org/10.1016/j.asmart.2023.09.002.

Hoffman, Hunter G., et al. "Virtual Reality Hand Therapy: A New Tool for Nonopioid Analgesia for Acute Procedural Pain, Hand Rehabilitation, and VR Embodiment Therapy for Phantom Limb Pain." Journal of Hand Therapy, vol. 33, no. 2, Apr. 2020, pp. 254–62. www.jhandtherapy.org, https://doi.org/10.1016/j.jht.2020.04.001.

Kang, Jae Myeong, et al. "Effect of Cognitive Training in Fully Immersive Virtual Reality on Visuospatial Function and Frontal-Occipital Functional Connectivity in Predementia: Randomized Controlled Trial." Journal of Medical Internet Research, vol. 23, no. 5, May 2021, p. e24526. www.jmir.org, https://doi.org/10.2196/24526.

Kikkert, Sanne, et al. "Neural Basis of Induced Phantom Limb Pain Relief." Annals of Neurology, vol. 85, no. 1, Jan. 2019, pp. 59–73. PubMed, https://doi.org/10.1002/ana.25371.

Kilteni, Konstantina, et al. "The Sense of Embodiment in Virtual Reality." Presence: Teleoperators and Virtual Environments, vol. 21, no. 4, Nov. 2012, pp. 373–87. Silverchair, https://doi.org/10.1162/PRES_a_00124.

Kwan, Rick Yiu Cho, et al. "Effects of Virtual Reality Motor-Cognitive Training for Older People With Cognitive Frailty: Multicentered Randomized Controlled Trial." Journal of Medical Internet Research, vol. 26, Sept. 2024, p. e57809. PubMed, https://doi.org/10.2196/57809.

Liu, Qian, et al. "Virtual Reality Technology in the Detection of Mild Cognitive Impairment: A Systematic Review and Meta-Analysis." Ageing Research Reviews, vol. 87, June 2023, p. 101889. ScienceDirect, https://doi.org/10.1016/j.arr.2023.101889.

Matamala-Gomez, Marta, Ana M. Diaz Gonzalez, et al. "Decreasing Pain Ratings in Chronic Arm Pain Through Changing a Virtual Body: Different Strategies for Different Pain Types." The Journal of Pain, vol. 20, no. 6, June 2019, pp. 685–97. PubMed, https://doi.org/10.1016/j.jpain.2018.12.001.

Matamala-Gomez, Marta, Tony Donegan, et al. "Immersive Virtual Reality and Virtual Embodiment for Pain Relief." Frontiers in Human Neuroscience, vol. 13, Aug. 2019. Frontiers, https://doi.org/10.3389/fnhum.2019.00279.

Matamala-Gomez, Marta, Mel Slater, et al. "Impact of Virtual Embodiment and Exercises on Functional Ability and Range of Motion in Orthopedic Rehabilitation." Scientific Reports, vol. 12, no. 1, Mar. 2022, p. 5046. DOI.org (Crossref), https://doi.org/10.1038/s41598-022-08917-3.

Matamala-Gomez, Marta, Francesca Frisco, et al. “Virtual Body Continuity during Action Observation Affects Motor Cortical Excitability.” Scientific Reports, vol. 15, no. 1, Apr. 2025, p. 13364. DOI.org (Crossref), https://doi.org/10.1038/s41598-025-97695-9.

Mateo, Sébastien, et al. “Motor Imagery Reinforces Brain Compensation of Reach-to-Grasp Movement after Cervical Spinal Cord Injury.” Frontiers in Behavioral Neuroscience, vol. 9, Sept. 2015. DOI.org (Crossref), https://doi.org/10.3389/fnbeh.2015.00234.

Qian, Jiali, et al. “The Effectiveness of Virtual Reality Exercise on Individual’s Physiological, Psychological and Rehabilitative Outcomes: A Systematic Review.” International Journal of Environmental Research and Public Health, vol. 17, no. 11, 11, Jan. 2020, p. 4133. www.mdpi.com, https://doi.org/10.3390/ijerph17114133.

Rábago, Christopher A., and Jason M. Wilken. “Application of a Mild Traumatic Brain Injury Rehabilitation Program in a Virtual Realty Environment: A Case Study.” Journal of Neurologic Physical Therapy, vol. 35, no. 4, Dec. 2011, pp. 185–93. DOI.org (Crossref), https://doi.org/10.1097/NPT.0b013e318235d7e6.

Ramachandran, V. S., and D. Rogers-Ramachandran. “Synaesthesia in Phantom Limbs Induced with Mirrors.” Proceedings of the Royal Society of London. Series B: Biological Sciences, vol. 263, no. 1369, Jan. 1997, pp. 377–86. royalsocietypublishing.org (Atypon), https://doi.org/10.1098/rspb.1996.0058.

Saby, Adam, et al. “Effects of Embodiment in Virtual Reality for Treatment of Chronic Pain: Pilot Open-Label Study.” JMIR Formative Research, vol. 8, no. 1, Feb. 2024, p. e34162. formative.jmir.org, https://doi.org/10.2196/34162.

Schuermans, Joke, et al. “Extended Reality in Musculoskeletal Rehabilitation and Injury Prevention - A Systematic Review.” Physical Therapy in Sport, vol. 55, May 2022, pp. 229–40. DOI.org (Crossref), https://doi.org/10.1016/j.ptsp.2022.04.011.

Sokołowska, Beata. “Impact of Virtual Reality Cognitive and Motor Exercises on Brain Health.” International Journal of Environmental Research and Public Health, vol. 20, no. 5, 5, Jan. 2023, p. 4150. www.mdpi.com, https://doi.org/10.3390/ijerph20054150.

Świdrak, Justyna, et al. “Virtual Embodiment in Fibromyalgia.” Scientific Reports, vol. 13, no. 1, July 2023, p. 10719. www.nature.com, https://doi.org/10.1038/s41598-023-36861-3.

Trujillo, Michael S., et al. “<p>Embodiment in Virtual Reality for the Treatment of Chronic Low Back Pain: A Case Series</P>.” Journal of Pain Research, vol. 13, Nov. 2020, pp. 3131–37. www.dovepress.com, https://doi.org/10.2147/JPR.S275312.

Moral Circles: Insights from Volunteer Medical Missions By Grace Peña

Abstract

Although recorded interviews with individuals on medical missions have indicated the benefits of such missions for all people involved, medical missions are not without their unintended consequences and risks, as the literature demonstrates. My experience participating in medical missions has inspired me to question the psychological forces that motivate volunteers to join missions, especially when they are helping people beyond the scope of their usual moral concern. This essay investigates the psychological concepts of moral circles and moral expansiveness, as well as present preliminary findings from a study in progress that is using the Moral Expansiveness Short Form, to determine the moral expansiveness of volunteers on medical missions.

Introduction

In the 21st century, the true effects of medical missions in underdeveloped nations have been questioned. There are several benefits of leading missions, as reported by most volunteers interviewed using questionnaire surveys. Most have expressed good experiences and opportunities to solidify key virtues such as compassion, cooperation, and self-sacrifice, along with the continued pursuit of medicine and the massive amount of patients seen. However, potential drawbacks exist as well, including the use of local populations as objects of study and as possibilities to gain personal experience, which partly strips developing countries of their self-reliance and promotes competition for local jobs.

My personal journey into this field of moral psychology began during a volunteer experience in Piura, Perú. The journey inspired me to perform a literature review of what psychological forces are behind altruistic actions, such as joining missions, and to explore what influences people to show concern for others not in their immediate moral circles. This essay also explores the evidence supporting and opposing medical missions and reviews the psychological concepts of moral circles and moral expansiveness. It also presents some of the preliminary results of a novel study in progress that is using the Moral Expansiveness Scale, Short Form (MESx), to assess the moral expansiveness of mission volunteers.

Volunteerism

Medical mission volunteers have long been portrayed as examples of altruism: people who sacrifice their time, resources, energy, and share their expertise to help those in suffering or in extreme need. True volunteers do not expect secondary gain in return, at least not in an economical or material sense (**Musick, 2008**). In a study in Brazil, 286 medical students were asked open-ended questions about their experience as a volunteer during the COVID-19 pandemic. The researchers concluded that 37% of them participated for altruistic reasons and only 19% had academic motivations (**Siqueira, et al., 2022**). In another study from the United Kingdom involving 38 interviews and 224 online surveys, most medical mission volunteers

reported altruistic motivations. The researchers also discovered that volunteers had very favorable opinions about their missions and expressed beliefs that volunteering can improve the quality of healthcare providers. They also reported beliefs that volunteer health care providers and state-sponsored medical volunteering benefit both the local community and the participants. Volunteers viewed the experience as a moment to learn from different health care systems outside the United Kingdom and to develop valuable skills. In addition, they were motivated by the need to feel useful and by the supposed impact of volunteer work (**McCauley, et al., 2021**).

The expression "service to humanity" is widely accepted as a common reason for pursuing a career in medicine in many countries **(Woodward, et al., 2017)**. Participating in medical missions appears to be associated with that expression, which aligns with the romantic dimension of being a doctor. According to an ethical analysis **(Onyebeke, et al, 2021)**, there are at least five professional virtues involved in constructing a medical mission: compassion, integrity, humility, self-effacement, and self-sacrifice. The most important virtue is likely compassion, meaning the ability to respond to the pain and suffering of others—in medical cases, that means patients. Indeed, the capacity to feel morally obliged to respond to others in suffering is characterized as being a key virtue among doctors.

It seems that for the volunteers, volunteerism is perceived as a paradigm of social service and defines what it means to be a doctor. Moreover, the people served seemed to share the same perspective. Community members have reported feeling hopeful as well as relieved that their problems had been acknowledged, and the medical assistance was a great help to the community in general. Beyond that, locals believe that health care professionals help with transferring knowledge to the community, which was testified in the Operation Smile missions in Colombia **(Martiniuk, et al., 2012)**.

However, having reviewed several articles related to medical volunteerism, I find that most studies have solely been collections of opinions from participants without much hard objective data. Indeed, data on the health outcomes of the populations served and the impact of volunteer missions by socioeconomic status remain unavailable. In one systematic review, which began with 7,400 articles from PubMed and EBSCO (i.e., the largest indexes for scientific literature in health care and the social sciences) that were narrowed down to only 41, the authors sought to clarify the psychosocial, diplomatic, and economic impacts of short-term medical missions. They concluded that although the most effective articles were the ones based on guidelines, most articles gathered were published with descriptive and reflective results. They added that the lack of data and analysis regarding short-term medical missions, their impacts, and their outcomes presumably stems from the common belief that missions are invariably altruistic and justifiable. Their article also reveals the lack of information about the economic costs and benefits of short-term medical missions, much less information about other outcomes and health or economic indicators of any kind (**Caldron, et al., 2015**). Even so, it seems that reporting the short-term effects of medical missions is more common than reporting their long-term effects. According to the authors of the mentioned systematic review, most reported studies have not met the parameters of evidence-based medicine. Such poor evidence and an over-reliance on personal

statements, experiences, and opinions can mask the real accomplishments of research in the field, while surveys have intrinsic limitations for their results and, in turn, obscure the true impact of missions trips (**Sykes, 2014**). For that reason, additional research is needed to strengthen the current evidence in order to accurately measure the effectiveness of volunteer missions.

Although medical missions offer numerous advantages, they are not free from criticism. For one, not everybody believes that volunteer missions are necessarily good and always achieve beneficial outcomes. On that count, the perceived success of medical missions typically hinges on the belief that developing countries need and will accept help from foreigners. For another, it is possible that international volunteering in developing nations takes away job opportunities in those countries, their self-reliance, and opportunities to improve their current conditions **(Bauer, 2017)**. The concentration on the benefits for volunteers, focus that should be on the situations of the people the mission is supposedly helping, demonstrate the true people the mission is targeted for. It is indeed conceivable that volunteer missions are created for secondary gains or, through their outcomes and/or poor organization, become dysfunctional and disruptive to the population that they seek to help.

In response, general guidelines have been proposed to facilitate the success of medical missions on local terms and to prevent collateral damage. Therefore, any proposed medical mission has to meet the four requirements in order to be launched. First, the highest-quality service to patients has to be a priority, as well as follow-ups with patients and supporting medical education. Second, the mission also needs to have a specific objective, though the ultimate goal should be to help the developing community become independent. Third, physicians need to project professionalism through the quality of care received by patients and through their acknowledgement of patients' autonomy. Fourth and finally, the safety of all participants in the mission has to be accounted for, which can be accomplished by coordinating travel, vaccinations, lodging, and so on (**Wilson, et al., 2012**).

Thus far, I have discussed volunteerism and how virtues such as altruism can motivate individuals to join volunteer medical missions. *Altruism*, defined as a selfless concern for the well-being of others, is often central to cultural ideals regarding doctors. However, genuine altruism can be difficult to maintain as modern medicine becomes increasingly complex, fragmented, episodic, and impersonal **(Jones, 2002)**, especially when both the healthcare system and society operate within narrow moral circles. In this paper, I investigate the relatively new concept of the moral circle and how it may influence volunteer activities. To illustrate that connection, I begin by sharing my personal experience pertaining to the topic.

Personal Reflections

My personal experience informed how I came across the topic of moral circles in the first place. I participated in my first Catholic medical mission in 2024, when I was 16 years old, along with my mother and brother. As a Peruvian American who wanted to see more of Perú and the true conditions of the country, I happily agreed to my mother's proposal to volunteer on a mission to Piura, Perú.

With a population of 33 million people, Perú achieved the status of an upper–middle–income economy in recent two decades according to the World Bank. From 2002 to 2013, the poverty rate fell from 60% to 24%, and from 2003 to 2024, GDP per capita jumped from US $2,126 in 2003 to $8,458 **(World Bank, 2024)**. However, despite its growing economy, Perú remains a country of obvious contrasts. For example, Perú performed very poorly during the COVID-19 pandemic as the country with the highest mortality rate in the world **(Ramos, et. al., 2022)**.

The specific medical mission that I joined in 2024 was affiliated with a church in Texas that has traveled to Piura once a year for several decades. The missionary organization, meanwhile, is firmly established in Perú and in the United States and organizes several mission trips every year. Due to the strong ties with the priest at the Santisimo Sacramento Catholic Parish in Piura, church groups from many states other than Texas, including Tennessee and Oklahoma, visit Piura as well. The parish there has developed a clinic over the past several years and, when available, provides free access to a psychiatrist, internal medicine, an OB/GYN office, an operating room, a dietician, and physical therapy.

The trip was eye-opening, to say the least. In a single week, I helped to deliver food packages and build houses, and I shadowed my mother in the clinic. I saw huge families living in bamboo homes with dirt floors, walking without shoes, and owning only what their sponsor families delivered. As mentioned, Perú is a country of stark contrasts and great disparity. Piura is a large state in the country's northeast, with a coastline on the Pacific Ocean and diverse geography encompassing both deserts and fertile valleys. What was striking to me in Piura was how the villages were not far—only a short drive away—from the downtown area. Even so, and despite it being the 21st century, villagers still had problems obtaining potable water and electricity, as well as eliminating and treating sewage, and their health care system remains small and unable to meet people's needs. Having so little, the villagers were immensely grateful for the bags of food, clothes, shoes and attention given by the mission's volunteers.

After several days of volunteering in the village, I spent a day with my mother in the mobile clinic. The first patient that I saw with my mother, a 12-year-old girl who underwent abdominal surgery several years ago, presented with a complaint of discomfort in her breast. I could not believe the challenges the young girl had confronted, much less the ones that lay in front of her. We could not do much for her, and that patient's story stuck with me.

Inspired by the kindness and the growth in faith of the villagers, I wondered whether the other missionaries felt as I did. Observing Piura's community and meeting so many families deeply impacted me, and I became concerned for their well-being. Being Peruvian American, I viewed each of those families as what mine could have been., in a parallel universe where I did not live in my pleasant, sheltered hometown. I came to care not only about my family and town but also for the people of Piura, and the following year, I found myself compelled to return.

Just a few months ago, I indeed revisited Piura. I was genuinely invested in dedicating myself to the people of Piura and willing to give up the comforts of my home, even if only for a week. During my second visit, I had opportunities to shadow physicians in the clinic and

operating room, to interpret for the clinic and pharmacy and three physicians specializing in family medicine, emergency medicine, and internal medicine, respectively. We traveled to places such as a neighborhood called " Los Polvorines," a hospice center founded by the parish, and chapels spread throughout different villages. Between the three physicians in the clinic, 775 patients were seen, and more than 25 patients were operated on by the surgical team. The work was exhausting. I observed physicians devote much of their time to coordinating and performing clerical work. In addition to seeing patients, the medical team stayed up late into the night to replace all the used medication—a time consuming, sometimes confusing process—while simultaneously interpreting each patient's history, filling out the prescriptions for more than 150 people a day, and following up on referrals. As I interpreted for the emergency physician, I understood the pain and helplessness that came with the position; even though we had established a clinic and were seeing as many people as we could, we could not help everyone. I remember a particular patient, a young mother of two, who presented with insomnia. When asked why she could not fall asleep, she opened her heart and told us of her worries and grievances at home. The physicians could not do much except send a referral to the psychiatrist, and all I could do was sympathize with her.

One day in particular was especially memorable for me. During a group reflection session, the mission leader asked every volunteer and provider to speak about their experience and how they felt. Each of us had to say one word to define our experience on the mission, and among the responses were "hopeful," "appreciated," "grateful," "faith," "redemption," "love," "refreshing," "eye-opening," and "compassion." Even though the work was frustrating and tiring, I saw that everyone felt rewarded in the end. It became clear to me that hard work and feeling tired at the end of the day were not reasons for frustration in the case at hand. On the contrary, they were sources of renewal.

The Moral Circle

While flying back to the United States, I asked myself, *What motivates people to leave their comfortable lives in cities across the country? Why do they come to poor, developing villages that have nothing to do with them?* I voiced those questions to my father, and he replied that the answer likely existed in the fields of psychology and sociology, which involve probing and explaining human behaviors, including volunteerism, altruism, and egoism.

Soon before my family's second trip, my father had given me *Moral Circles* by Jeff Sebo. It was through Sebo's book that I was introduced to the underlying philosophy and ideas about moral circles, and I thought about how impactful medical missions truly are. I began to wonder whether my two experiences had broadened my own moral circles.

At first, it was difficult for me to understand the concept of the moral circle. Is it a psychological phenomenon, or a social process? What psychological forces or cognitive forces play a role in creating or evaluating moral circles? Is it an ethical or moral concept? Many people have tried to explain the concept, beginning with Peter Singer, who Sebo quotes heavily in his book. Although Singer explains a great deal related to social behaviors, he offers no

straightforward definition of *moral circle*. However, he does cite the explanation of W. E. H. Lecky in the first pages of his book **(Singer, 2011):**

> The question of morals must always be a question of proportion or of degree. At one time the benevolent affections embrace merely the family, soon the circle expanding includes first a class, then a nation, then a coalition of nations, then all humanity, and finally, its influence is felt in the dealings of man with the animal world. In each of these stages a standard is formed, different from that of the preceding stage, but in each case the same tendency is recognised as virtue. **(Lecky, 1869)**

More than 140 years after Lecky, Steven Pinker wrote about moral circles in his famous book, *The Better Angels of Our Nature: Why Violence Has Declined* as a sphere in which every being is considered morally significant. Pinker explains how humanity has increased the moral circle by reducing violence throughout the ages **(Pinker, 2012)**.

Published this past year, Jeff Sebo in his book *The Moral Circle* details the ethics and philosophy behind moral circles. He defines the *moral circle* as the group of beings who matter for their own sakes. For humans, there may not be a particular reason why certain sentient beings matter. Sebo states that every member of the moral circle should matter equally, but there is confusion about what beings belong in the moral circle **(Sebo, 2025)**. History has had a tremendous influence over that debate.

After reading articles in the psychological and sociological literature, I now better understand what the moral circle is. Simply put, it is an outcome of all the mental processes in a person's brain. All the psychological forces therein—love, hate, altruism, egoism, cooperation, ideology, empathy, and personal experience—lead each of us to decide and define which individuals are deserving of our moral obligation and ethical treatment. The moral circle is a personal, psychosocial outcome representing only the specific individual. As a product of social pressure and ideology, the moral circle varies not only between individuals but also between cultures and societies. The moral circle also varies in intensity, for we can feel and act more energetically in the service of people whom we consider to be of major moral concern, which is usually our immediate family members and friends.

At the same time, the moral circle is a relatively late outcome and is likely formed as we mature into adults, when we reach maturity. Research has shown that children have larger moral circles than adults, but those results may be due to an instinctive feeling, not quite the moral circle itself **(Marshall, et al., 2025)**.

Understanding the moral circle as an outcome makes it easier to realize how complex and dependent it is upon many other processes, as mentioned above. By extension, moral expansiveness is a psychosociological construct that can be understood as a model for investigating the moral circles of individuals in order to explain the forces that urge humans to produce moral circles in the first place. The model proposes a validated instrument that is relatively easy to use in surveys. Although the moral circle and moral expansiveness are

different, they are related as a referent and reference. Finally, the moral circle can be investigated from a perspective or in light of an ethical or philosophical ideal. Sebo, Pinker, and Singer, by comparison, understand it as a moral or ethical dimension and discuss which circles are best for humanity or most societies. From a prescriptive way they ultimately conclude that a wide one—a sphere encircling all sentient beings and individuals is probably the best for humanity.

The Psychology of Moral Expansiveness and Volunteerism: Preliminary Results

The models that seek to represent and explain moral circles have evolved. Initially, moral circles were evaluated following three approaches that eventually evolved to the one that uses the Moral Circle Expansiveness scale (MES). In 2016, the MES was created to determine the span and depth of moral concern. Along those lines, **Crimston et. al. 2016 and 2018** have proposed 30 entities, each given a grade when placed in a ring of the moral circle (see Figure 1). The range of the moral circle is determined according to the placement of 30 entities, and the depth is evaluated according to where the entities are placed and the cost of the willingness to defend them. The aggregate score compiles the values of the entities in each ring of the circle and represents the expansiveness of the individual, as shown in Figure 2**.** The 30 entities cover 10 categories: family and friends, in-group, out-group, revered individuals, marginalized individuals, villains, high-sentience animals, low-sentience animals, plants, and environment. Scores are given according to the intensity within the various moral circles; the inner circle of moral concern is worth 3 points, the outer circle of moral concern is worth 2, the fringes of moral concern are worth 1, and the area outside moral concern is worth 0. Thus, the minimum score for any individual is 0 points, whereas the maximum score is 90. (**Crimston et. al. 2016)**

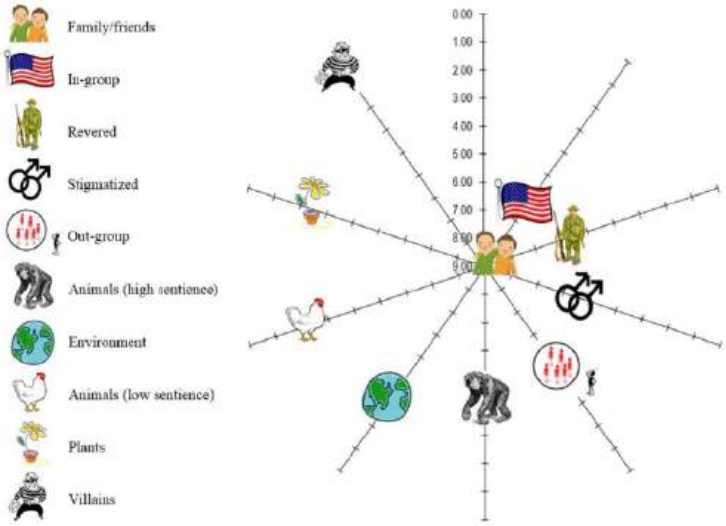


Fig 1: The 10 categories of entities included in the MES are demonstrated above, with an example of the entities in a moral circle. Graph from Crimston, C.R., et al. "Toward a Psychology of Moral Expansiveness." *Current Directions in Psychological Science*, vol. 27, no. 1, 2018, pp. 14-19.



Fig 2: The order of the rings of the moral circle is portrayed above. Image from Crimston, C.R., et al. "Moral Expansiveness: Examining Variability in the Extension of the Moral World." *Journal of Personality and Social Psychology,* vol. 111, no. 4, 2016: pp. 636-53.

In 2018, the Moral Expansiveness Scale, Short Form (MESx), with 10 entities from each core category, was developed to reduce participant burden and improve efficiency when distributing the form. The removal of the other 20 entities allows participants to complete the survey faster and prevents the influence of politically divisive entities. The resulting maximum score is 30 points, instead of 90. When the MES and MESx were compared the MESx was found to be suitable for measuring the aggregate score, it maintains high reliability and validity. **(Crimston, et al., 2018)**.

I am currently involved in a research study **(Peña, et al. 2025)** designed to determine whether people who volunteer for medical missions have higher scores on the MESx than the general population. It also seeks to pinpoint whether participating in a medical mission significantly affects the aggregate score on the MESx, compared with scores before participating, for volunteers who are participating in a volunteer mission for the first time. Above all, the study seeks to gauge whether or not participating in a medical mission influences the growth of an individual's moral circle. Along with administering the MESx to every participant, we also obtained their demographic data and asked them whether they had previously volunteered for a mission.

In the study, convenience sampling was used to recruit participants, naming by contacting various volunteer mission organizations that travel to Perú from the United States. We also received authorization from the Peruvian American Medical Society and the Santisimo Sacramento Catholic Parish Missions in Piura. Next, we sent the MESx pre-mission and post-mission surveys in an online survey developed using Google Docs to missionaries traveling to Piura and Iquitos, Perú, in June and July 2025. From the group that traveled to Piura, we

received 25 pre-mission responses.

The mean MESx score was 19.47 points for individuals who had participated in a mission before and 21.6 points for first-time participants. By comparison, from the group that traveled to Iquitos, we received 16 pre-mission responses; the mean score was 18.10 points for experienced individuals and 15.67 points for first-time participants. As for the post-mission survey results, the Iquitos group earned a mean of 18.50 points. All respondents to the post-mission survey from Iquitos were first-time participants, which indicates an increase in their mean MESx aggregate score.

Although these results remain preliminary, the mission volunteers scored higher than the average in the United States. On that count, the mean MES score is estimated to be 44 points in the southern United States and 45 points in the northern United States **(Kirkland, et al., 2023).** Converted to MESx scores, those values are equivalent to approximately 14.67 and 15 points, respectively. As future missions occur, more data will continue to be gathered. Our aim is to survey at least 250 participants, which will likely take us an entire year.

To this day, not a single published study has attempted to measure moral expansiveness using MES scores in a group of volunteers on medical missions. As a consequence, it remains unclear whether such behavior or the proclivity to sign up for volunteer missions correlates positively with a deeper and/or broader moral circle. It is also unknown whether participating in a volunteer mission causes an individual to broaden their moral circle and whether it correlates with more altruistic actions or prosocial attitudes or activities.

Several studies on external motivational factors have shown what can influence people's moral circles. If we face a society or situation in which resources are limited and/or individuals compete for them, it is possible for our moral circle to shrink (**Crimston, et al. 2018)**. It seems that the ideology we adopt also influences our moral circles, for significant differences exist between the moral circles of conservatives and liberals **(Waytz, et al., 2019)**. Social factors influence our moral circles as well. As shown in another study (**Kirkland, et al., 2023)**, great variation also exists in moral circles across countries (n = 36) due to varying levels of trust and economic inequality.

Although researchers have rarely used the MES or MESx to identify events that can affect an individual's moral expansiveness, evidence suggests that some key experiences can broaden a person's moral circle. In that vein, scholars have evaluated whether an experience of awe can influence growth in moral expansiveness (**Song, et al., 2023**), given that awe has been observed to impact cognition, emotion, and behavior. The method in that study involved different presentations of awe using videos, slow-motion videos, and a simulation. Added to experiencing awe, participants completed the MESx. Although the results revealed no significant difference in moral expansiveness in light of experience of awe through videos, self-reports from participants indicated an increase in moral expansiveness nevertheless. Following awe demonstrated through virtual reality, participants indicated increased moral expansiveness.

Another study involving an event possibly influencing moral circles showed an increase in moral expansiveness following a compassion-focused therapy seminar (**Kirby, et al., 2024**).

That seminar, lasting 2 hours, aims to improve intrapersonal and interpersonal well-being. Two weeks after the seminar, there were increases in the aggregate score towards family and friends and revered individuals. Three months afterward, there was an increase in all 10 categories of entities.

Beyond that, another study marks the first to ever evaluate changes in moral circles due to psychedelic experiences (**Olteanu & Moreton, 2025**). That study involved surveying participants who had had psychedelic experiences about how they acted before and after the experience. Using the MESx to measure the participants' moral expansiveness before and after, the researchers calculated mean MESx scores before and after of 19.42 and 23.01, respectively, which indicates a significant increase in moral expansiveness.

Those three studies are the only instance of research that I have found so far that has evaluated the effectiveness of a life-changing event on an individual's moral expansiveness. An event that should be considered along those lines, similar to experiences of awe and educational interventions, is participating in a medical mission. The experience of seeing people who are suffering and/or living in starkly different, even disadvantageous circumstances, and the experience of being unable to provide full care can probably ignite a person's compassion and/or activate emotional centers in the brain that motivate the expansion of their moral circle. The study that I am helping to develop is pursuing that possibility by administering the MESx before and after such missions, in the hope of providing the aggregate scores and other details of people who are volunteering for the first time.

Conclusion

Medical missions are an outstanding opportunity to show one's commitment to helping people in need. Based on studies that have involved asking volunteers about their experiences on medical missions, it seems that the experience is a gratifying one and possibly an opportunity to invest in several virtues, including compassion and self-sacrifice. Although the scientific evidence about their true impact in the populations and countries served remains scarce, research on the volunteers has shown that altruism is likely the chief motivator for them. Even so, how volunteers perceive their moral circles, meaning the group of individuals who are believed and felt to be deserving of moral obligation and ethical treatment, may also play a role. It is possible that people with deeper and broader moral circles are inclined to engage in more prosocial activities, one being to volunteer and another being to provide medical care. In any case, more research is needed to fully answer those questions, and, to achieve that task, the psychosociological model offered by moral expansiveness theory using the MESx provides a promising platform.

Works Cited

Bauer, I. More harm than good? The Questionable Ethics of Medical Volunteering and International Student Placements. Tropical Disease Travel Medicine Vaccines, vol. 3, no. 5, 2017, https://doi.org/10.1186/s40794-017-0048-y

Caldron, P.H., et al. "A Systematic Review of Social, Economic and Diplomatic Aspects of Short-Term Medical Missions." BMC Health Services Research, vol. 15, no. 1, 2015, https://doi.org/10.1186/s12913-015-0980-3.

Crimston, C.R., et al. "Moral Expansiveness Short Form: Validity and Reliability of the MESx." PLOS ONE, vol. 13, no. 10, 2018, p. E0205373, https://doi.org/10.1371/journal.pone.0205373.

Crimston, C.R., et al. "Moral Expansiveness: Examining Variability in the Extension of the Moral World." Journal of Personality and Social Psychology, vol. 111, no. 4, 2016: pp. 636-53, https://doi:10.1037/pspp0000086

Crimston, C.R., et al. "Toward a Psychology of Moral Expansiveness." Current Directions in Psychological Science, vol. 27, no. 1, 2018, pp. 14-19.

Jones, R. "Declining Altruism in Medicine." British Medical Journal, vol. 324, no. 7338, 2002, pp. 624–625, https://doi.org/10.1136/bmj.324.7338.624.

Kirby, J.N., et al. "A Brief Compassion Focused Therapy Intervention Can Increase Moral Expansiveness: A Randomized Controlled Trial." Mindfulness, vol. 15, 2024, pp. 282-294. https://doi.org/10.1007/s12671-023-02300-7.

Kirkland, K., et al. "Moral Expansiveness around the World: The Role of Societal Factors across 36 Countries." Social Psychological and Personality Science, vol. 14, no. 3, 2022, pp. 305–318, https://doi.org/10.1177/19485506221101767.

Lecky, W. History of European Morals from Augustus to Charlemagne, vol. 1. D. Appleton, 1869. Online Library of Liberty, pp 101. https://oll.libertyfund.org/titles/lecky-history-of-european-morals-from-augustus-to-charlemagne-vol-1.

Marshall J. et al. "When Development Constricts our Moral Circle". Nature Human Behaviour. 2025 May 28;9635. Epub 2025 May 28. doi: 10.1038/s41562-025-02212-7

Martiniuk, A. L., et al. "Brain Gains: A Literature Review of Medical Missions to Low and Middle-Income Countries." BMC Health Services Research, vol. 12, no. 1, 2012, https://doi.org/10.1186/1472-6963-12-134.

McCauley, M., et al. "Experiences and Impact of International Medical Volunteering: A Multi-Country Mixed Methods Study." British Medical Journal Open, vol. 11, no. 3, 2021, p. e041599, bmjopen.bmj.com/content/11/3/e041599, https://doi.org/10.1136/bmjopen-2020-041599.

Musick, M.A., Wilson J. Volunteers : A Social Profile. Bloomington, Ind., Indiana Univ. Press, 2008.

Olteanu, W., Moreton S.G. “Meaningful Psychedelic Experiences Predict Increased Moral Expansiveness.” Journal of Psychoactive Drugs, 12 Mar. 2025, pp. 1–9, https://doi.org/10.1080/02791072.2025.2478095.

Onyebeke, W. et al. “A Professional Virtues-based Ethical Framework for Medical Missions.” American Journal of Obstetrics & Gynecology Global Reports vol. 1, 4 100017. 30 July 2021, doi:10.1016/j.xagr.2021.100017

Pinker, S. The Better Angels of Our Nature : Why Violence Has Declined. Penguin Publishing Group, 2011.

Peña G., et al. “Research Project Proposal: Measuring Moral Expansiveness Before and After a Voluntary Medical Mission in Peru”, 2025. (Upon request)

Ramos W., et al."Epidemiological Characteristics of Deaths from COVID-19 in Peru during the Initial Pandemic Response".Healthcare, vol. 10, no. 12: 2404, 2022, https://doi.org/10.3390/healthcare10122404

Sebo, J. The Moral Circle: Who Matters, What Matters, and Why. W. W. Norton & Company, 2025, pp 3.

Singer, P. The Expanding Circle. Princeton University Press, 2011. (pp. 1)

Siqueira, M.A.M., et al. "Medical Students’ Participation in the Volunteering Program During the COVID-19 Pandemic: a Qualitative Study about Motivation and the Development of New Competencies”. BMC Medical Education, vol. 22, no. 111, 2022. https://doi.org/10.1186/s12909-022-03147-7

Song, J.Y., et al. “Awe Promotes Moral Expansiveness via the Small-Self.” Frontiers in Psychology, vol. 14, No. 6, 2023, https://doi.org/10.3389/fpsyg.2023.1097627.

Sykes, K.J. “Short-Term Medical Service Trips: A Systematic Review of the Evidence.” American Journal of Public Health, vol. 104, no. 7, 2014, pp. e38–e48, https://doi.org/10.2105/ajph.2014.301983.

World Bank. “The World Bank in Peru.” World Bank, 5 Apr. 2024. Accessed July 8 2025. https://www.worldbank.org/en/country/peru/overview.

Waytz, A., et al. “Ideological Differences in the Expanse of the Moral Circle.” Nature Communications, vol. 10, no. 1, 2019, pp. 1–12, https://doi.org/10.1038/s41467-019-12227-0.

Wilson, J. W., et al. “Rules of Engagement: The Principles of Underserved Global Health Volunteerism.” The American Journal of Medicine, vol. 125, no. 6, 2012, pp. 612–617, https://doi.org/10.1016/j.amjmed.2012.01.008.

Woodward, A., et al. “Reasons to Pursue a Career in Medicine: A Qualitative Study in Sierra Leone.” Global Health Research and Policy, vol. 2, no. 1, 2017, https://doi.org/10.1186/s41256-017-0054-7.

Reinventing the Beat: Syncopation as the Heart of American Musical Innovation
By Timothy Yang

Abstract

Syncopation, the intentional disruption of rhythmic expectations, has played a central role in the evolution of American music. This paper explores the historical and cultural trajectory of syncopation, from its origins in African polyrhythms to its defining influence on ragtime, jazz, rock, funk, and hip-hop. Drawing on research from musicology, cognitive neuroscience, and cultural history, the study demonstrates that syncopation does more than challenge rhythmic norms. It activates the brain's predictive and reward systems, producing a blend of anticipation and resolution that deepens emotional engagement and drives creative innovation. Through a comparative analysis of Jon Batiste's reinterpretation of Beethoven's *Für Elise*, the paper shows how rhythmic complexity can transform the familiar into something strikingly new. The case highlights a broader pattern in American music, where syncopation operates not merely as a stylistic device but as a structural force behind reinvention. Supported by empirical findings on listener enjoyment and neural activation, the argument advances a central claim that syncopation functions as a cognitive mechanism for novelty, a cultural tool for transformation, and a rhythmic thread connecting the past to the present. It is in the unexpected beat that American music has found its most enduring voice.

Introduction

"Syncopation is the lifeblood of jazz, the pulse of funk, the unpredictability of hip-hop."
—Miles Davis, Jazz Legend
"A composer's duty is to surprise—anything else is mere routine."
—Igor Stravinsky, Composer

The interplay between anticipation and resolution has been a driving force behind the evolution of music. In American popular music, syncopation, a technique that disrupts expected rhythmic patterns, has repeatedly pushed boundaries and redefined genres. The original syncopation techniques, such as cross-rhythms and off-beat accents, can be traced to African rhythmic systems (Anku 212). These rhythms inspired ragtime pioneers like Scott Joplin, who wrote the syncopated piano melodies that made dance halls explode with energy. By the early 20th century, syncopation became a defining feature of jazz (Gioia, *History of Jazz* 35).

In the mid-century, syncopation's evolution continued through rock and roll, with Chuck Berry's backbeat and syncopated guitar riffs propelling the genre to mainstream success. Motown's signature sound and funk's percussive bass lines brought syncopation to new heights, while hip-hop turned rhythmic disruption into an art form, with offbeat rhymes and beats reimagining musical structure (Temperley, "Origins" 24; Witek et al. 2). Across generations, syncopation has become a favored tool for creating the signature sounds of American music, reflecting changes in cultural influences, technologies, and aesthetic priorities.

Yet this evolution is not random. It is reflective of an underlying psychological necessity. Music triggers dopamine release in the brain, particularly during the anticipation phase, when listeners are waiting for their favorite part of a song (Salimpoor et al. 259). Over time, songs "get old" as the brain adapts to predictable rewards. To remain fresh and engaging, music must evolve, and thus, musicians must innovate. This interplay between expectation and disruption not only captivates listeners but also fosters innovation, as artists build upon historical precedents to redefine musical boundaries (Huron, *Sweet Anticipation* 110).

American music has evolved through the transformative power of syncopation, which has consistently disrupted rhythmic expectations to redefine genres and inspire creativity. Rooted in African polyrhythms, syncopation became central to the development of ragtime, jazz, rock, and hip-hop. It has acted as a unifying thread in the evolution of American musical identity.

Syncopation is an effective tool for innovation because of its cognitive impact. By engaging the brain's reward pathways and creating anticipatory dynamics, syncopation fosters emotional engagement and enhances the listener's experience (Vuust and Frith 599–600). This ability to captivate attention and drive creative breakthroughs in both composition and performance has positioned syncopation as a key force in musical evolution.

Jon Batiste's reimagining of Beethoven's *Für Elise* illustrates how syncopation can transform a familiar classic. His version introduces rhythmic complexity that deepens the emotional impact and enhances the overall listening experience. By increasing the use of syncopation, Batiste creates a rendition that feels both innovative and compelling, which in turn heightens enjoyment. This is confirmed by research and performer insights which reveal how the rhythmic shifts challenge and inspire creativity. Batiste's *Für Elise* demonstrates syncopation's power to reinvent the familiar, showing how it keeps American music dynamic and transformative.

The Evolution of American Music

If you had wandered into Congo Square on a Sunday afternoon in the late 19th-century New Orleans, you would have witnessed something extraordinary. Enslaved Africans, granted rare moments to gather, drummed, danced, and sang with rhythms that seemed to pulse with life itself. These were not just sounds. They were statements of defiance, survival, tradition, and connection. Syncopation, the unexpected shift of rhythm, was at the core of these performances, a technique that disrupted predictability and drew listeners into the energy of the moment (Gushee 4). This rhythmic rebellion did not fade. It grew into a force that would shape American music. Its emergence also marked a pivotal moment in American cultural history, where the fusion of African and European traditions birthed a rhythmically unique identity that continues to evolve.

Scholars agree that syncopation emerged as a prominent musical element in America at the turn of the 20th century, coinciding with the emergence of ragtime (Temperley, "Origins" 19). Up until then, American music had predominantly followed European musical tradition. And from that point, syncopation became the lifeblood of American music. It began as many

transformative ideas do, against adversity. African slaves brought to the United States against their will, brought with them the mastery of a rich and complex library of rhythms that had been sharpened over dozens of generations. These rhythms combined with European harmonic structures to create something wholly new. For example, on cotton plantations, spirituals emerged, blending African call and response patterns with European hymnals. Later, the blues extended this further with 12-bar structures and syncopated riffs that bent and stretched time itself. In addition, vocal phrasing in blues often drags or anticipates the beat, reflecting African rhythmic elasticity (Gioia, *History of Jazz* 21–23). While African-Americans kept their traditions alive as a form of resistance and solace, White America appropriated them. Minstrel shows, a very popular form of entertainment in the 19th century, took the rhythms and stories of African culture and turned them into caricature (Burns and Ward 47). And yet, even during this injustice, the seeds of something revolutionary were planted. Syncopation began to creep into mainstream music, changing it forever.

At the heart of this revolution was New Orleans. The city was the perfect petri dish to spawn a new American sound. It was a place where the rules bent and broke in ways that did not happen anywhere else in America. Peter Latrobe, an architect from New York, described it poignantly when he wrote, "Everything has an odd look," and it was "impossible not to stare at a sight wholly new, even to one who has traveled much in Europe and America" (qtd. in Bern 14). By the mid-19th century, it was the most cosmopolitan city in America, a place where African, European, Caribbean, Native American, and Asian cultures collided. It boasted 3 opera houses, a Creole orchestra, and a White orchestra. (Burns and Ward, 8) By 1860, 40 percent of the people of New Orleans were foreign born. (Burns and Ward 6) Contrary to popular belief at the time, slaves were often highly talented musicians themselves. Tourists "were sometimes startled to hear among the cries of street vendors the sound of slaves singing arias" (Burns and Ward 9). Congo Square became a sanctuary where slaves could dance to rhythms that echoed their ancestral memory. These dances were known for their rhythmic complexity, polyrhythms, and syncopation. Fascinated and sometimes bewildered, some visitors called it "authentic African music." But it was not just African. It was more than that.

There in the heart of New Orleans, these beats eventually turned into a new form of American music, ragtime, a progenitor of jazz. In the early days of New Orleans, the line between jazz and ragtime was so fine that the two terms were used interchangeably, and it is only in scholarly hindsight we can distinguish the two genres (Gioia, *How to Listen to Jazz* 18). Ragtime introduced a dynamic interplay between steady bass lines and syncopated melodies, fusing African rhythms with European harmonic traditions. Scott Joplin's "Maple Leaf Rag" exemplified this innovation, showcasing a vibrant new style that reflected a cultural transformation (Gioia, *History of Jazz* 35). This melding of diverse influences not only shaped ragtime and jazz but also laid the groundwork for the rhythmic language of most subsequent American popular music. And the numbers tell the story. By 1910, African Americans formed 30% of the male musician and music teachers in the United States, doubling from the 1900 census (Gioia, *History of Jazz* 38). The demand for a syncopated style had taken root.

Critics of the time could not ignore the syncopation that made ragtime so distinct. This rhythm, unpredictable and fresh, became the hallmark of the genre. But where did it come from? Some traced its roots to the British "Scotch snap," a rhythmic quirk where a sixteenth note precedes a dotted eighth (Temperley, "Syncopation in Rock" 20). Others pointed to a far more compelling origin. It was the rich rhythmic traditions of Africa first showcased to American audiences during minstrel shows in the Antebellum South (Anku 221–224). These rhythms laid the foundation for what would become one of the defining characteristics of American music.

As African Americans migrated north during the Great Migration in pursuit of better wages and treatment in northern cities, they carried jazz with them to cities like Chicago and New York. This Great Migration carried the hopes, dreams, musical styles, and syncopation of the descendants of slaves to the North and West, forever reshaping the American musical consciousness. The embrace of African rhythmic style in American music made U.S. music much more open to innovation. At the heart of it, syncopation drove much of the musical innovation that formed jazz, rock and roll, bebop, blues, and other American genres. The statistics bear this out. The average number and density of syncopated events in American music increased in a linear fashion from the 1890s to the 1930s, with every decade incorporating more events than the last (Huron and Ommen 218–219).

Jazz's signature swing feel came from syncopation, where the second and fourth beats of a measure, known as backbeats, were emphasized, and eighth notes were unevenly subdivided (Nelias et al.). This rhythmic phenomenon defined America's popular music throughout the 1920s and 30s. Louis Armstrong's "West End Blues" epitomized the swing. Syncopated trumpet improvisations and virtuosic piano passages introduce unexpected accents and rhythms, breaking from strict metrical timing and surprising listeners (Gioia, *History of Jazz* 57). In jazz, the syncopation interacts with European-style chord progressions, allowing soloists to improvise melodically while following the harmonic changes.

However, swing music had its critics. A growing segment of African-American pioneers grew frustrated with the commercially successful yet rhythmically simple genre. Like the earlier minstrel shows, critics argued that swing music was a form of cultural appropriation. White musicians often got the better gigs and commanded higher fees than equally or even more talented African-Americans. George Gershwin's famed jazz concerto "Rhapsody in Blue" was emblematic of this controversy, as African-American artists, such as Duke Ellington, felt it overshadowed and even plagiarized their work and rhythms (Iverson). Another complaint was that swing compositions such as the Rhapsody simplified the rich African rhythmic tradition to a level even white musicians could understand. One particularly egregious example of this occurred when pianist Oscar Levant sold many recordings of his performance of the Rhapsody that demonstrated his inability to grasp even Gershwin's basic rhythms (Goldberg). Coincidentally, several twentieth-century shrewd white music executives forced African-American musicians to sign one-sided contracts that did not allow them to get their fair share. African-American pioneers, frustrated by the appropriation and oversimplification of their art, sought a new path forward.

This frustration gave rise to bebop in the 1940s. Bebop emerged as a direct and indignant reaction to the injustices in the jazz world. It represented a complex form of jazz, less accessible to the mainstream. This style, pioneered by musicians like Charlie "Bird" Parker in late-night Harlem jam sessions, would come to be known as bebop (Meadows 124). It featured blisteringly fast tempos that defied the danceable grooves of swing, syncopations that twisted and turned in ways that made even seasoned listeners lose their footing, and a near-Herculean level of technical and rhythmic mastery. Given the motivation and context, this level of complexity certainly did not appear accidental. Pioneers of bebop designed their music as a fortress, too intricate to be easily imitated or appropriated.

The arrival of bebop marked a pivotal moment in America's evolving relationship with syncopation. Before bebop, the recipe for popular music was straightforward: the more syncopated the rhythm, the greater the appeal. From ragtime to swing, syncopation was a seductive novelty that brought listeners to their feet. But bebop flipped the formula. Its relentless complexity overwhelmed audiences rather than inviting them in. This phenomenon is rooted in cognitive overload, the psychological reasons for which will be discussed later in the paper (Huron, *Sweet Anticipation* 148). For the first time, jazz had stopped trying to charm the public. Bebop was not designed for a broad audience. It was crafted for the artist, the virtuoso, the few who could keep up with its breakneck pace and intricate melodic turns.

For a time, bebop felt revolutionary. But as its tempos soared and its melodies grew denser, it alienated casual listeners. Few dared dance to it. As we will see later in the paper, bebop was too fast, too complex, and too cerebral for people to process. This complexity would ultimately sever jazz's connection to the everyday audience, leaving a vacuum in popular music. That absence would demand something new to fill it.

Into that opening stepped rock and roll. Emerging in the late 1940s and early 1950s, rock simplified syncopation again, making it accessible while retaining its rebellious energy. Rock harnessed new technologies such as the electric twang of a guitar, the backbeat of a drum kit, and the exuberant swagger of a new generation.

As bebop rose, jazz lost share to rock, which became the preferred musical genre of America's youth beginning in the late 1940s. Advancements in radio, the emergence of a large teenage population with significant spending power, and the start of the civil rights movement naturally gave rise to a cosmopolitan, socially conscious genre. A fusion of blues and gospel with country music and swing, rock symbolized teenage rebellion against the rigid social constructs imposed by traditional parents. Rock also defied systemic racism through its integration into the mass media. To the dismay of many parents, young rock fans listened to and attended concerts of stars of all races.

Rock had its own characteristic rhythms, but they were heavily based on existing art forms. As Dr. Steven William of the University of Southern Indiana summarizes, "If you look at basic rock 'n' roll, the fundamental formula is basically African American blues with a little more speed and electricity. Then you add bass and drums, and suddenly you've got something new" (Williams).

Unlike the complex and improvisational syncopation found in bebop, rock and roll simplified syncopation to make it more accessible. The genre relied on repetitive yet incessant syncopated patterns, creating a steady pulse that encouraged movement and engagement. Jazz's emphasis on the backbeat and use of offbeat accents played an especially vital role in shaping the sound and energy of the genre (Temperley, "Syncopation in Rock" 32).

New technology, such as electric guitars and drum kits, also magnified the impact of syncopation to give it a raw and more powerful edge. Elvis Presley's "Hound Dog" demonstrates this approach, using the Latin tresillo guitar riff and the 12-bar blues form to drive the beat forward. This syncopation also became central to rock and roll's identity, symbolizing its youthful rebellion. The off-beat nature of syncopation mirrored the genre's challenge to the rigid, predictable rhythms of earlier styles.

Despite their structural simplicity, rock rhythms continued to elicit criticism from segments of the American public, particularly among those shaped by early twentieth-century racial ideologies. As Christopher John Farley observes, white audiences were often more receptive to rock music when it was performed by white artists such as Elvis Presley, despite the genre's foundational roots in African American musical traditions. This disparity reflects a broader social anxiety among older generations—many of whom matured during the era of the Ku Klux Klan—that musical forms influenced by Black culture might encourage interracial socialization or relationships. Nevertheless, public efforts to suppress rock 'n' roll only intensified its appeal, as its dynamic energy and cross-cultural resonance fueled widespread popularity (Farley).

Rock's rhythms are exemplified in Chuck Berry's "Johnny B. Goode," where a syncopated guitar riff pairs with a strong backbeat, channeling the rhythmic vitality of both blues and jazz. Additionally, early rock emphasized off-beat accents in guitar and piano parts, creating a sense of unpredictability and groove. Little Richard's "Tutti Frutti" is a prime example, with its syncopated piano rhythms driving the song's infectious energy.

The development of syncopation in American popular music has played a foundational role in shaping modern genres such as pop, soul, and rap. In his article *The Origins of Syncopation in American Popular Music*, David Temperley traces how rhythmic displacement—particularly second- and fourth-position syncopation—originated in 19th-century African American musical traditions and evolved through ragtime, jazz, and rock into contemporary styles (Temperley 18–19). He explains that fourth-position syncopation, in particular, creates a strong anticipatory effect that disrupts metric expectations and adds expressive energy to a piece (Temperley 23–24). Temperley also highlights the influence of African and Caribbean rhythmic structures, such as cross-rhythm and the tresillo pattern, which have been absorbed into mainstream Western music, thereby enriching its rhythmic complexity (Temperley 34–36). This lineage is evident in contemporary tracks like Billie Eilish's "Bad Guy," where syncopated vocal lines emphasize weak beats and anticipate strong ones, reinforcing the ongoing impact of historical rhythmic practices in shaping modern musical sensibilities.

Soul music leverages syncopation to enhance groove and emotional expression, building on the genre's roots in gospel and rhythm and blues. As noted by *Music Rising* at Tulane University, soul developed in the American South by fusing the emotional intensity of gospel with the structure of rhythm and blues, employing techniques such as melisma, call and response, and gospel-inspired harmonic progressions ("Soul"). These elements, combined with rhythmic innovations like triplets and the shift from shuffle to even-note grooves, laid the groundwork for soul's signature sound. Aretha Franklin's "Respect" exemplifies this, with off-beat horn stabs and gospel-influenced vocal phrasing that create a commanding groove. Stevie Wonder's "Superstition" showcases syncopation through its iconic keyboard riff, interacting with the bass and drums to form a tight, funk-infused rhythm. These examples demonstrate how soul artists have used syncopation not just as a rhythmic tool, but as a core expressive device rooted in African American musical traditions.

In hip-hop, syncopation is central to both lyric delivery and beat production. As pitch is deemphasized, syncopated lyrics alone keep listeners engaged. Rappers use syncopation to craft intricate flows that play with rhythmic timing, while producers layer beats with off-beat accents and unexpected breaks. Kendrick Lamar's "HUMBLE." highlights this, with his use of irregular phrasing and pauses. This vocal syncopation creates a sense of unpredictability and tension.

Syncopation also facilitates modern-day genre blending through danceable rhythms. Beyoncé's "Formation" merges pop, soul, and rap with syncopated trap-inspired beats and layered vocal rhythms, creating a powerful statement of identity and culture. Missy Elliott's "Get Ur Freak On" combines syncopated beats, inspired by an Indian percussion instrument known as a tabla, with off-beat vocal phrasing, blending global rhythmic influences into a hip-hop context. Genre blending facilitates the creation of new sounds of music as well as new beats to dance to, giving us a glimpse into how American music might continue to evolve.

From jazz and rock, syncopation began a journey that continues to shape American music today. It found new expressions in soul, R&B, and hip-hop as changing culture and new technology helped form these genres. Artists used syncopation to craft beats that were both complex and accessible, with melodies that lingered in memory and lyrics that moved with rhythmic surprise. Over time, it became a driving force in American music from pop, rap, and electronic music. Through every cultural shift and technological innovation, syncopation remains the constant, the spark that continues to drive the innovation into new musical genres.

The Power of Syncopation

But why is syncopation such a driving force? At its core, the pleasure of syncopated rhythms may be connected to how our brain's auditory location systems process temporal patterns and spatial cues. It challenges and rewards our brain's predictive mechanisms that help us make sense of the world around us by exploiting our auditory systems' sensitivity to location and time, our predictive processing, and our reward mechanic for cognitive effort (Huron, *Sweet Anticipation* 109–110).

First, our auditory system is highly sensitive to temporal patterns because it evolved to help us locate sounds in the environment. This is biologically useful for many reasons, including the location of possible prey and predators. This involves processing subtle differences in the timing, and intensity of sounds arriving at each ear. It's here that a concept called Neural Darwinism comes into play. Our brain does not just passively process the world but it models different views of the world, which constantly compete. New dominant models adapt and evolve the brain's circuits, reinforcing the neural pathways that best predict and respond to incoming sensory information. Syncopation, in this sense, is like a workout for the brain. It disrupts predictable timing, forcing the brain to engage more attentively with the rhythm to track where the "beat" is. For instance, in Chuck Berry's "Johnny B. Goode," the syncopated guitar riff interrupts the steady backbeat, creating rhythmic surprises that engage the brain and drive the song's infectious energy.

This heightened engagement can be pleasurable, as the brain resolves the tension created by these temporal irregularities. Specifically, as the brain's auditory cortex constantly anticipates when and where a sound will occur based on patterns. Lower-level brain regions estimate predictions of expected input based on information about one's surroundings from higher-level areas. This generates a recursive process, which tries to minimize the difference between the prediction and the event that occurs. If it is accurate, the brain triggers activity in reward-related regions like the dorsal and ventral striatum and dopaminergic pathways, which are associated with pleasure (Salimpoor et al. 259–261). Using positron emission tomography (PET) scanning, scientists have been able to directly correlate the dopamine release to anticipatory and experiential pleasures when listening to music.

Syncopated rhythms violate these expectations in subtle and manageable ways, creating a blend of surprise and reward. In fact, empirical studies have measured cognitive impact of syncopated rhythms versus non-syncopated rhythms. When a melody moved from a syncopated rhythm to an unsyncopated rhythm, the syncopated patterns were rated happier than their non-syncopated counterparts (Keller and Schubert 145). And when researchers measured syncopated patterns in a non-serial fashion, they found that syncopated rhythms were enjoyed more than unsyncopated ones (Witek et al. 3–4). This brain's reward system explains why syncopation felt so revolutionary in ragtime's heyday. Scott Joplin's *Maple Leaf Rag* didn't just entertain. It sparked a kind of neurological thrill, the anticipation that neuroscientists now link to dopamine release. Researchers have shown that "…melodies that violate rhythmic expectations were rated as more enjoyable and 'happier' than rhythmically predictable melodies" (Vuust and Frith 599). And as audiences craved more of this rhythmic excitement, composers answered by pushing boundaries and crafting entirely new genres in their quest to surprise and delight. However, the correlation between unpredictability and enjoyment is not linear. Once rhythms cross a threshold of complexity the brain's auditory cortex cannot accurately predict where a sound will be. It gets alarmed and overwhelmed, and instinctively prepares for an incoming predator attack. In essence, overly complex music ceases to be enjoyable, as we have seen with bebop.

Beyond these mental rewards, syncopation taps into something deeper, our primal urge to move. Syncopation and rhythm, more broadly, do not just stimulate the mind, they demand a physical response. Research shows that the more people feel compelled to dance to a piece of music, the more they enjoy it, suggesting a direct link between syncopation and movement (Witek et al. 4). While the exact nature of what it is about music that offers a pleasurable desire to move is unclear, we can logically conjecture that these rhythms can also create the perception of "motion" in music, as listeners track shifting accents and anticipate where the next beat will land. This experience engages spatial attention systems, which are tied to auditory localization. As with the auditory localization mechanism we describe above, the movement activates similar dopamine reward pathways that encourage us to stay attuned to irregular auditory patterns.

And this presents itself in the form of dance. Syncopation fulfills our natural desire to dance to music. It has been shown that the more people want to dance to music, the more they enjoy it (Witek et al. 2). Interestingly, this trend is universal to all humans, even infants. Melodies which violate rhythmic expectation are more enjoyable and happier than predictable melodies, proving syncopation increases lilt. In the act of dancing, or even foot-tapping, to groove-based music, the body coordinates its movements to be beat-directed and periodic. As previously discussed, the additional pleasure elicited by danceable rhythms helps explain why syncopation frequently serves as a unifying element in genre blending. This phenomenon also sheds light on the decline in bebop's popular appeal. As its tempos accelerated, the music became increasingly inhospitable to dancing, thereby severing the physical connection that had long been integral to the listener's musical experience.

To summarize, syncopated rhythms need to strike a balance between predictability and complexity, making them engaging but not overwhelming. When the music becomes overwhelming, the music becomes less accessible. This balance is critical to enjoyment. If it is too simple, rhythms align with expected beats and do not stimulate attention. However, if it is too complex, rhythms become chaotic, and the brain cannot form a coherent prediction, let alone dance to the music. Rhythms with moderate syncopation are pleasurable because they provide just enough unpredictability to keep the auditory system actively engaged (Keller and Schubert 149).

Even if you find the right balance, syncopation's effect is not a stable phenomenon because the brain adapts to the predictable patterns and rewards within a piece of music. The initial excitement fades, and the song begins to feel repetitive or even dull. The excitement diminishes exponentially with repeated exposures. This process, known as habituation, explains the need for musical innovation and the genre formation we discussed in American music (Rankin et al. 135–136). Composers and songwriters needed to continuously push boundaries, introducing new rhythms, textures, and harmonic surprises to reignite the brain's engagement and stave off predictability.

While habituation drives the need for composers to innovate music beyond the familiar, syncopation is a potent method to be able to evolve the music. In America, syncopation became more than a technique. It became the creative force behind the birth of jazz, the rebellion of rock,

and the energy of hip hop. Syncopation disrupts the expected, forcing listeners to engage with the music in fresh and exciting ways. It has shaped the very identity of American music.

A case study that compares a classical European piece to one that contains significant syncopation, but with the same core baseline melody, is useful in illustrating syncopation's transformative power. We examine Jon Batiste's reimagining of Ludwig Von Beethoven's *Für Elise*, where syncopation breathes new life into a familiar classic.

The Case Study of Jon Baptiste's Für Elise: Syncopation as a Driver of Innovation

Jon Batiste's reimagining of Beethoven's *Für Elise* illustrates how much syncopation can elevate enjoyment of a piece. A score analysis of both the Beethoven and Batiste versions was conducted to assess their respective degrees of syncopation. Since various definitions exist, we define the widely used Longuet-Higgins & Lee definition as core syncopation: "Syncopation is a note in a beat followed by a rest or continuation on the next beat, where the second beat is stronger than the first" (Temperley, "Origins" 19–21). Five additional types of devices capable of producing a syncopated effect were also measured, thereby broadening the operational definition of syncopation.

The analysis revealed striking differences in the use of rhythmic techniques. Beethoven's composition demonstrates a restrained approach, with no instances of core syncopation. Beethoven's works demonstrate a restrained approach. It had no instances of core syncopations. When using the expanded definition, there were one hemiola, two uses of ornamentation, one fermata, no meter changes, and ten tempo changes, totaling just fourteen occurrences. In contrast, Batiste's interpretation is rhythmically rich and complex. It features 58 core syncopations, 22 hemiolas, 63 instances of ornamentation, 8 fermatas, 16 meter changes, and 16 tempo changes, totaling 183 occurrences.

Table 1.1: The number of rhythmic devices used in Beethoven and Batiste versions

Version	Core Syncopations	Hemiolas	Ornamentation (including those written out)	Fermata	Meter Change	Tempo Change	**Total**
Beethoven	0	1	2	1	0	10	**14**
Batiste	58	22	63	8	16	16	**183**

Although the definitions employed in this study are reasonably specific, further clarification is warranted. In traditional Western Classical performance practice, the first beat of each measure typically carries the strongest emphasis. Depending on the genre, the second and third beats may vary in strength—for instance, the sarabande emphasizes the second beat over

the third, while the waltz gives prominence to the third beat. For the purposes of this analysis, we treat the second and third beats as rhythmically equivalent. Thus, a note held from the second beat into the third is not considered syncopation.

If one note is tied through multiple stronger beats, it is marked as only one instance of syncopation. If the melodic contour is unclear (for instance, one melody turns into two melodies or is split between the clefs), no addition is made to the syncopation tally. Moreover, slurs in Beethoven's version not printed in Batiste's version are not counted as syncopation. This ensures fairness across both editions. For instance, Beethoven's version has several two-note slurs where the first note is a weak beat and the second note lands on a strong beat. According to the Western Classical tradition, this gesture is considered syncopation because, in a two-note slur, the first note is generally emphasized. However, these slurs often do not occur in Batiste's version.

As you can see, the results are clear. Batiste's version relies heavily on syncopation, while the original score exhibits no core syncopations. When the expanded definition is used, Batiste uses over 13 times the number of syncopations that Beethoven uses. Beethoven's original is rhythmically structured, relying on repetition and predictability rather than tension and release. Batiste introduces unexpected accents, off-beat rhythms, and irregular phrasing that continuously disrupt predictable patterns.

These rhythmic innovations result in a listening experience that feels alive, unpredictable, and emotionally charged. Rather than leaning on the familiarity of a well-worn melody, Batiste uses syncopation to reframe and recontextualize the piece, offering listeners fresh engagement with something they thought they knew. The cognitive tension introduced by the rhythmic displacement creates a sense of anticipation and surprise, which are hallmarks of musical pleasure and innovation as discussed in earlier sections.

From the performer's perspective, the unexpected syncopations and jazz-inspired runs make the piece much more technically challenging than Beethoven's original version. But the syncopation heightens anticipation by shifting accents off the expected beats and introducing pauses or irregularities that keep listeners guessing. Performers are challenged to navigate the interplay between structure and improvisation, embodying the tension and release that the audience experiences. This dynamic approach requires a heightened sense of musical intuition and creativity, reinforcing the notion that syncopation drives not only listener engagement but also the performer's artistic exploration.

In both design and execution, Batiste's work illustrates that syncopation is not merely a stylistic flourish, but a tool of transformation. It turns familiarity into novelty, structure into expression, and heritage into evolution. In doing so, it echoes the broader history of American music, where syncopation has consistently served as the engine of reinvention. This interpretive framework is further supported by empirical research, which demonstrates that syncopation significantly enhances listener enjoyment and engagement compared to rhythmically predictable compositions. (Witek et al. 4; Keller and Schubert 145).

Conclusion

Syncopation is more than a musical technique. It is a mechanism for innovation. From the defiant rhythms of enslaved Africans in Congo Square to the intricate flows of Kendrick Lamar, syncopation has been central to American music, challenging conventions and driving creativity. By disrupting predictable rhythms, it balances familiarity with surprise, reflecting the tension between stability and change that underpins artistic progress. This rhythmic disruption has redefined genres and remains a core force shaping musical evolution and cultural identity (Temperley, "Origins" 34; Gioia, *History of Jazz* 39).

It is too simple, however, to characterize syncopation's impact as random artistic innovation. We have an underlying biological need for change. The human brain craves patterns. It loves predictability, rewards it, and builds entire stories around it. And yet, nothing captures our attention like the unexpected accent. Something that is in a place where there should not be one. A pause that stretches just a little too long. Syncopation pulls us into this space between order and chaos, where the predictable is disrupted just enough to make us lean in, wondering what comes next. This effect is what makes syncopation so potent. It resonates deeply with our cognitive wiring, engaging the brain's reward systems and inviting us to move, feel, and think in new ways (Salimpoor et al. 261; Vuust and Frith 599).

Yet, it is not static. Syncopation must constantly evolve, as the human brain quickly adapts to predictability. This drives the need for change (Rankin et al. 135–136). Jon Batiste's reimagining of *Für Elise* encapsulates this dynamic beautifully. By weaving syncopation into the familiar, he transforms a classical composition into a modern masterpiece, reminding us that innovation often arises not from abandoning tradition but by reinterpreting it. His work exemplifies how syncopation serves as both a bridge between the past and the future and a spark that ignites cultural transformation.

Looking forward, syncopation's significance extends far beyond music. It offers a blueprint for understanding how creativity thrives under constraints, how the tension between comfort and surprise fosters engagement, and how even the most established ideas can be reinvented. Syncopation teaches us that disruption is not merely an act of rebellion but a form of dialogue. It is a way to challenge boundaries while honoring the systems we seek to change. As music continues to globalize, syncopation will undoubtedly play a vital role in blending traditions, creating hybrids, and pushing genres to new frontiers. But its essence will remain the same. It is the heartbeat of innovation, a reminder that the unexpected is where we find not only rhythm but meaning.

By unpacking the mechanics and magic of syncopation, we gain more than an understanding of music. We uncover a framework for exploring how humans create, adapt, and connect. In its unpredictable beats lies the pulse of progress. It is a rhythm that ensures the song of humanity never stops evolving.

Works Cited

Anku, Willie. “Principles of Rhythm Integration in African Drumming.” Black Music Research Journal, vol. 17, no. 2, 1997, pp. 211–238.

Bern, Medea Isphording. Images of America: San Francisco Jazz. 7th ed., Arcadia Publishing, 2014.

Burns, Ken, and Geoffrey C. Ward. Jazz: A History of America’s Music. 7th ed., Alfred A. Knopf, 2000.

Cohn, Richard. “A Platonic Model of Funky Rhythms.” Music Theory Online, vol. 22, no. 2, June 2016, https://mtosmt.org/issues/mto.16.22.2/mto.16.22.2.cohn.html.

Davis, Miles. Quoted by Timothy Yang. Reinventing the Beat: Syncopation as the Heart of American Musical Innovation. Unpublished manuscript, 2024.

Farley, Christopher John. “Elvis Rocks. But He’s Not the First.” TIME, 6 July 2004, https://content.time.com/time/magazine/article/0,9171,664995,00.html. Accessed 9 July 2025.

Gioia, Ted. How to Listen to Jazz. Hachette Book Group, 2017.

---. The History of Jazz. 3rd ed., Oxford University Press, 2021.

Goldberg, Jacob. “Local 802 During the Swing Era.” Allegro, vol. 117, no. 2, Feb. 2017, https://www.local802afm.org/allegro/articles/local-802-during-the-swing-era/.

Gushee, Lawrence. “The Nineteenth-Century Origins of Jazz.” Black Music Research Journal, vol. 14, no. 1, Spring 1994, pp. 1–24.

Huron, David. Sweet Anticipation: Music and the Psychology of Expectation. MIT Press, 2006.

Huron, David, and Ann Ommen. “An Empirical Study of Syncopation in American Popular Music, 1890–1939.” Music Theory Spectrum, vol. 28, no. 2, Fall 2006, pp. 211–231.

Iverson, Ethan. “The Worst Masterpiece: ‘Rhapsody in Blue’ at 100.” The New York Times, 26 Jan. 2024, https://www.nytimes.com/2024/01/26/arts/music/george-gershwin-rhapsody-in-blue.html.

Keller, Peter E., and Emery Schubert. “Cognitive and Affective Judgements of Syncopated Musical Themes.” Advances in Cognitive Psychology, vol. 7, Dec. 2011, pp. 142–156.

Meadows, Eddie S. Jazz Scholarship and Pedagogy. 3rd ed., Taylor & Francis Group, 2006.

Nelias, Corentin, Eva Marit Strum, Thorsten Albrecht, York Hagmayer, and Theo Geisel. “Downbeat Delays Are a Key Component of Swing in Jazz.” Communications Physics, vol. 5, no. 237, 2022, https://doi.org/10.1038/s42005-022-00995-z.

Quillen, Daniel Josh. “Review of ‘Anatomically Distinct Dopamine Release During Anticipation and Experience of Peak Emotion to Music’ (Salimpoor et al.).” Neuroanthropology Class Blog, University of Alabama, 24 Apr. 2019, https://dehb.ua.edu/neuthroanthropology-class-blog/review-of-anatomically-distinct-dopamine-release-during-anticipation-and-experience-of-peak-emotion-to-music-salimpoor-et-al.

Rankin, Catherine H., et al. “Habituation Revisited: An Updated and Revised Description of the Behavioral Characteristics of Habituation.” Neurobiology of Learning and Memory, vol. 92, no. 2, Sept. 2009, pp. 135–138.

"Soul." Music Rising: The Musical Cultures of the Gulf South, Tulane University, https://musicrising.tulane.edu/discover/music-genres/soul/. Accessed 9 July 2025.
Salimpoor, Valorie N., et al. "Anatomically Distinct Dopamine Release During Anticipation and Experience of Peak Emotion to Music." Nature Neuroscience, vol. 14, Feb. 2011, pp. 257–262.
Stravinsky, Igor. Quoted by Timothy Yang. Reinventing the Beat: Syncopation as the Heart of American Musical Innovation. Unpublished manuscript, 2024.
Tan, Ivan, Ethan Lustig, and David Temperley. "Anticipatory Syncopation in Rock: A Corpus Study." Music Perception, vol. 36, no. 4, 2019, pp. 353–370.
Temperley, David. "Syncopation in Rock." Popular Music, vol. 18, no. 1, 1999, pp. 19–40.
---. "The Origins of Syncopation in American Popular Music." Popular Music, vol. 40, no. 1, 2021, pp. 18–41.
Vuust, Peter, and Chris D. Frith. "Anticipation Is the Key to Understanding Music and the Effects of Music on Emotion." The Behavioral and Brain Sciences, vol. 31, no. 5, Oct. 2008, pp. 599–600.
Williams, Steve. "Rock 'n' Roll and 'Moral Panics'—Part One: 1950s and 1960s." University of Southern Indiana News, 20 Feb. 2017, https://www.usi.edu/news/releases/2017/02/rock-n-roll-and-moral-panics-part-one-1950s-and-1960s.
Witek, Maria A. G., et al. "Syncopation, Body-Movement and Pleasure in Groove Music." PLOS One, Apr. 2014, pp. 1–12. https://doi.org/10.1371/journal.pone.0094446.