|CURIEUX|

ACADEMIC REVIEW

February Issue

Issue 4 Part 3

Editing Staff

Chief-In-Editor

George Hasnah

Lead Editor

Amanuel Geremew

Ashley Du

Assisting Editors

Jing Yan Huang

Nidhi Potineni

Sophia Kong

Tony Yuan

Ridit Aggarwal

Manu Sharma

Rishi Gadikota

Eric Wang

Table Of Contents

A Global Community: Three Theories, One Puzzle By Stellan Milford

Among scholars of international relations, three theories compete to best explain how states interact with each other and how international organizations fit in the puzzle. Realism, liberalism, and constructivism each work to encompass the seemingly infinite number of factors that construct the global order. At any given moment, one of the three theories may best suit the current circumstances, but circumstances change and glaring shortcomings appear. Scholars can be fixated on their favorite theory and fail to see how different theories can intertwine. For example, both realists and liberals appreciate the principles behind the strategy known as external balancing, when states form alliances to counter a threat, but approach the interaction from different directions. The realists view the scenario through the lens of power dynamics, while liberals see it as a demonstration of the benefits of cooperation. Adherence to only one theory does not account for the fact that the world is chaotic and unpredictable. Like nations and other actors, these theories and their advocates can seem stuck in their distinct cultures. For realists, anarchy brings fear, followed by the hardheaded response they see as natural or inevitable. Liberals hope progress and commerce — and, yes, the authority of international organizations — will smooth over our anarchic tendencies. Maybe we can understand anarchy in a positive sense, as is implicit in the constructivism of Alexander Wendt and others. There is room for creativity. Arguments from all three perspectives need to be pieced together to solve the global puzzle of international relations, and international organizations can play a role in making that happen.

The term "realism" in international relations theory refers to the view that self-interested states compete for power and security on an international stage. Under this umbrella, multiple schools of thought take different approaches to the baseline importance of power, the inevitable competition for more of it, and the nations who are by far the chief competitors—international organizations are a sideshow for realists. Since the late 20th century, the preferred approach has been structural realism. This perspective was preceded by classical realism, a theory advanced by Hans Morgenthau, among others. Classical realists point to human nature when explaining the centrality of power in international relations, assuming that it is human nature to crave dominance over other people and thus for states to try to dominate their neighbors and rivals. Classical realism thrived during the period of imperialism, when many power-hungry empires and states were all demanding a piece of the global pie. The situation was multipolar; power was distributed to three or more states. For example, when the Austro-Hungarian Empire, the Ottoman Empire, Germany, Russia, France and Great Britain jockeyed for position, the resulting conflicts led to WWI. These global powers conquered the smaller states for glory, resources, and land. The realist argument is this: there is no law or policy that can get in the way of the simple human desire for hegemony; from the classical point of view, war is inevitable.

Structural realism also assumes that conflict is unavoidable, but its proponents get to that "wisdom" by a different route. At its core, structural realism points to the distribution of power and, more specifically, the struggle for power that derives from the network of states and, more

broadly, the international system. Structural realism holds to a preconceived notion that every single state wishes to survive and thrive, but how they do this can differ in a multitude of ways: offensive strategies, defensive strategies, internal balancing, etc. This perspective focuses solely on states and their relations with others; considerations generated by international institutions or by religious and cultural ties are secondary at best ("*International Relations: One World, Many Theories*," *Foreign Policy*, no. 110, Stephen M. Walt, page 31).

Structural realism has different schools of thought. Offensive realism stems from the idea that to gain more power and to survive requires an aggressive attitude, often leading to war. One state cannot be certain about the intentions of another state even if they agree to terms. Not only this, but within offensive realism there is the appeal of preventative war: attacking a smaller developing nation that might surpass the power of your own country or pose a threat to your own country in the near future. Realists of this type cite the example of the Peloponnesian War, when the Spartans attacked Athens over worries that the larger and more commercial Athens was gaining influence at Sparta's expense. Much more recently, we saw another example of preemptive warfare when the United States conducted an air strike on Iranian uranium enrichment plants over fears that Iran was developing nuclear weapons.

While the offensive view focuses on the need to use force, defensive realism argues that a state builds up its own security forces to drastically raise the price other states would pay for attacking it. When state A sees state B building up their defensive capabilities, state A will also inherently begin to build up their security as well. In turn, every nation will continually build up their security together, essentially creating a world in which there are massive fortresses dotting the globe — but destructive wars are avoided (Walt, 37). The Cold War exemplified the principles of defensive realism: it was a bipolar era of power distribution, with the USSR and USA constantly increasing their own nuclear/military capabilities. This arms race created Mutually Assured Destruction or MAD, which kept both states from attacking each other.

Realists tend to take a cynical view about how international organizations —which are supposed to be impartial entities for the common good of the world — interact with states. The main crux of their argument is that IOs mirror the balance of power of the states that make them up. IOs are used as resources for a state actor's particular interests ("*The False Promise of International Institutions," International Security 19,* no. 3, John T. Mearsheimer, page 7). The most powerful organ of the United Nations, the UN Security Council, gives extraordinary power to the five permanent members: the USA, Russia, the United Kingdom, France, and China. With its policies binding for the 193 members of the UN, The UNSC is the most glaring example of how states use IOs for their self-benefit ("Theory, Methods, and International Organizations," *International Organizations: Politics, Law, Practice,* Ian Hurd, page 20-21). For example, as Russia was just beginning its invasion against Ukraine, Russia vetoed the policy that condemned its unjust invasion of Ukraine. To take a more historical example, the US used the power of the UNSC to bend the rules of engagement, allowing them to send their military into Korea in 1950 (*The United Nations and Changing World Politics,* Thomas G. Weiss, page 40). The realist

perspective centers around power, and IOs are viewed as just another way for a state to gain power.

For liberals, friendship rather than fear maintains international peace. Contrary to realism, liberalism takes a more optimistic view: cooperation and democracy promote the healthiest international relations. One important version of this liberal tradition was established by the philosopher Immanuel Kant. His admirers describe a Kantian triangle: the three keys to international peace are democracy, economic interdependence, and international organizations. The first corner of the triangle, democracy, is elaborated upon by Democratic Peace Theory (DPT), grounded in Kant's work and in the work of those inspired by him. Democratic nations, according to this theory, don't go to war with each other because the costs of war (both the money spent and the citizens killed or wounded) are not so easily tolerated as they are in nations with authoritarian governments. Politicians in democratic nations won't be reelected if they go to war because war is bad for business and civilian life ("Liberelism," *IR Theories,* Bruce Russet, page 82). Domestic policymakers are incentivized to make decisions that avoid war. Because of the mature republican institutions and the shared rationalism needed to sustain a democracy in the first place, Kant proposed that democratic states tend to seek peaceful resolutions by nature.[7]

The second side of the triangle, economic interdependence, also serves a deterrent to war. After WWII, world leaders noticed the connection between material happiness and peace and worked tirelessly to establish the economic network that we have (for the time being) today. States understand the penalties for leaving or challenging this network, which deters them from starting wars or worsening disputes, although the recent example of Russia-Ukraine and the sanctions that followed tell us that deterrence is never foolproof. In this age of globalization and interconnectedness, a war involving only two states can raise prices for commodities worldwide. These costs are felt long before a war is fought. Take the tariffs that the United States has levied on multiple goods and nations, for example; this one action by the US has caused mass disruption across the globe, imposing negative economic effects on billions.

The final piece to the liberal triangle of peace is international organizations and institutions. Unlike realists, liberals believe that IOs help facilitate peace, in part by operating independently from states. Ian Hurd separates IOs into three groups: actors, fora, and resources. IOs as actors behave like independent entities; prime examples are institutions like the International Court of Justice or the International Criminal Court. Such institutions keep states in check and put a cap on what they can and cannot do (Hurd, 17-18). IOs as fora (or, forums) focus more on managing interactions between states; this tool-like approach can take shape in many ways, from serving as a physical meeting place (like the UN headquarters in New York City) to acting as impartial referees for conventions, agreements, and treaties. IOs that are forums can solve disputes, hold states accountable for the treaties they sign, and reduce the misunderstandings between countries (Hurd, 19-20). From the liberal perspective, IOs as a resource means that these institutions help states fulfill their duties to their civilians and to the international community.[11] Such IOs help maintain and manage the economic interdependence

Kant posited is so crucial for peace. All of the liberal arguments intertwine in order to achieve peace.

Relatively new in the field of international relations theories, constructivism, more than realism or liberalism, respects how ideas and beliefs play out and shape international relations. Constructivists pay attention to culture as much as they do to basic considerations of power (the realist bias) or universal progress (the liberal hope). In the case of the US's recent strikes on Iranian nuclear development sites, constructivists might ask why the U.S. attacked Iran when there are many other states that already have nuclear weapons (e.g., the United Kingdom, France, Israel, etc.). It all boils down to the fact that the US is allied with the UK, France, and Israel but sees Iran as an enemy, reasons that have little to do with a balance of power. The US government doesn't believe that Iran is rational; the U.S. has the idea that Iran will always be inclined to use their nuclear weapons for evil purposes. From the constructivist perspective of Alexander Wendt and others, the anarchical world is what you make it out to be ("Social Constructivism," *Introduction to IR — Theories and Approaches*, Jackson Sorenson, page 162). This view leaves room for creativity, with constructivist arguments for why states both comply and don't comply with international organizations. A simple answer for why states comply with IOs is because they believe in the norm that the specific IO is encouraging (Hurd, 17).

International relations are not as bleak as realism makes them out to be. States tend to look for more power but not in the simple way that most realists propose. After the fall of the Soviet Union and the effective ending of the Cold War, the U.S. never aimed for or achieved the power anticipated by realists. Over the last 35 years, the world has remained relatively peaceful because of the points outlined in Kant's liberal triangle. We find ourselves in a situation where states are cooperating while aiming to maintain power, a position explained through both liberalism and realism. Yet, these perspectives on international organizations remain too specific, as if they are determined to oppose each other. In the manner of realists, some IOs, like the United Nations Security Council, represent the balance of power; in the manner of liberals, other IOs, like the International Court of Justice, do in fact act independently from states, operating as forums to hold states accountable. The constructivist view, albeit broader and more vague, better represents how states can interact with international organizations. If a country's leadership believes that an IO represents the state's national identity, then it will cooperate with this IO. Alternately, states will also sign onto IOs if they see their "friends" already within the organization. These institutions are what each state perceives them to be (Sorenson, 168). One can achieve a more comprehensive view of the international system by applying different perceptions from liberalism and realism, and constructivism offers that freedom.

Works Cited

Hurd, Ian. "Theory, Methods, and International Organizations." International Organizations: Politics, Law, Practice, November 25, 2010, 1-28. PDF.

Mearsheimer, John T. "The False Promise of International Institutions." International Security 19, no. 3 (1995): 5-49. PDF.

Russet, Bruce. "Liberalism." IR Theories, 68-86. PDF.

Sorenson, Jackson. "Social Constructivism." Introduction to IR ? Theories and Approaches, March 7, 2022, 162-77. PDF.

Walt, Stephen M. "International Relations: One World, Many Theories." Foreign Policy, no. 110 (Spring 1998): 29-46. PDF.

Weiss, Thomas G., David P. Forsythe, Roger A. Coate, and Kelly-Kate Pease. The United Nations and Changing World Politics. 8th ed. Routledge, 2019.

Evaluating the Impact of Symptom Selection on Machine Learning Models for Type 2 Diabetes Screening By Cyrus Darki

ABSTRACT

Type 2 diabetes (T2D) is a chronic metabolic disease that can cause long-term damage to multiple organs if not diagnosed and treated early. Because T2D often remains undiagnosed in its early stages and access to laboratory testing is limited in many settings, symptom-based screen approaches may help identify individuals at increased risk. This study evaluates the effectiveness of multiple machine learning (ML) models for predicting T2D using a small set of symptom-based clinical features. A publicly available Kaggle dataset containing 520 patient records was analyzed. Two experiments were conducted to assess the impact of symptom selection on predictive performance. Experiment 1 used three hallmark symptoms of diabetes–polyuria, polydipsia, and sudden weight loss–while Experiment 2 expanded the feature set by adding weakness and visual blurring. Five ML classifiers (Decision Tree, Random Forest, Logistic Regression, Neural Network, and K-Nearest Neighbors) were trained and evaluated, along with a majority voting ensemble. Model performance was assessed using accuracy, precision, recall, and F1 Score. Across both experiments, all models demonstrated strong and consistent performance with test accuracies ranging from 87% to 88%. Adding secondary symptoms did not lead to substantial improvements in predictive accuracy. These findings suggest that a small set of key symptoms captures much of the predictive signal for diabetes and that symptom-based ML models may serve as useful tools for diabetes screening, especially in resource-limited settings.

INTRODUCTION

Type 2 Diabetes (T2D) is one of the leading causes of death and disability worldwide. In 2021, an estimated 529 million people were living with diabetes, and researchers project that by 2050, more than 1.31 billion people will be affected (Article 12). T2D is a chronic metabolic disease caused by excess glucose in the blood (hyperglycemia), which can lead to vision loss, heart disease, stroke, and kidney failure (Article 13). Ideally, diagnosis is made with blood tests when patients are asymptomatic, but unfortunately, health care access is limited for many individuals, including those in low and middle-income countries (Article 23). As such, clarifying symptom-based risk factors, with the help of artificial intelligence and machine learning, can identify symptom combinations that can improve screening and treatment of diabetes in those with the greatest need.

The pathogenesis of T2D involves insulin resistance and pancreatic β-cell dysfunction leading to hyperglycemia (Article 14). When food is consumed, the body turns carbohydrates into glucose. With the help of insulin, a hormone secreted by the pancreatic β- cells, glucose enters the peripheral tissue and is used as energy. In T2D, insulin resistance in peripheral tissues causes more and more insulin to be required to get sufficient uptake of glucose in the cells. The β- cells initially compensate for the insulin resistance by secreting more insulin, but over time,

they fail to keep up and lose partial function (Article 14). The hyperglycemia during the course of diabetes leads to microvascular and macrovascular complications with long-term damage in several organs such as the eyes, kidneys, nerves, heart, and blood vessels (Article 7, Article 15).

Early diagnosis and treatment are crucial to help prevent complications, but continue to be a challenge in certain demographics and countries. In the United States, the proportion of undiagnosed diabetes has declined substantially over recent decades due to the use of screening blood tests (Article 16, Article 21, Article 22). Nonetheless, approximately 9.5%-17.8% of all diabetes cases were still left undiagnosed in 2017-2020 (Article 16). Undiagnosed diabetes was more prevalent in older and obese adults, minorities such as Asian Americans, and those without health care access, with rates ranging from 23 to 61% (Article 16). Moreover, globally, the International Diabetes Federation (IDF) estimated that in 2021, nearly one in two adults 20 to 79 years old with diabetes were unaware of their diabetes status, almost 239.7 million people (Article 17). The highest proportions of undiagnosed diabetes were found in Africa, the Western Pacific, and South-East Asia regions. The IDF concluded that diabetes surveillance needs to be strengthened to reduce the prevalence of undiagnosed diabetes mellitus in low- and middle-income countries (Article 17).

When routine health care is not accessible, symptom-based screening can help identify individuals at higher risk for undiagnosed diabetes and may serve as a component of diabetes risk assessment tools (Article 18). Hallmark symptoms of marked hyperglycemia, the defining clinical feature of diabetes, include polyuria (excessive urination), polydipsia (excessive thirst), and weight loss, while additional symptoms may include polyphagia (excessive hunger), blurred vision, impairment of growth, weakness, and susceptibility to certain infections (Article 7). Although certain symptom combinations can increase diabetes likelihood, more research is needed to clarify which symptom sets can improve identification (Article 19, Article 20).

Artificial intelligence (AI) and machine learning (ML) models are becoming increasingly applied to diabetes screening, diagnosis, and treatment (Article 6). AI and ML are powerful tools for performing predictive analysis and extracting meaningful knowledge from big datasets. Furthermore, AI and ML have the potential to facilitate rapid, low-cost diabetes screening, support high-risk populations, and improve access to care for individuals who face barriers to in-person medical evaluation (Article 2).

Several prior studies have investigated the use of ML for diabetes diagnosis. In one study using a dataset of 520 patients, eight symptoms contributed to diabetes risk when occurring together: gender, polyuria, polydipsia, sudden weight loss, weakness, blurred vision, partial paresis, and obesity. These symptoms co-occurring increased the chances of developing diabetes by 1.63 times (Article 1). In another study using three ML classification algorithms (Decision Tree, SVM, Naive Bayes) to identify diabetes in patients, it was determined that Native Bayes performed with the highest accuracy of 76.30% compared with other algorithms (Article 3). In a larger study with 6561 respondents, four classifiers (Naive Bayes, Decision Tree, Adaboost, and Random Forest) were applied, and a combined logistic regression and random forest-based model achieved an accuracy of 94.25% (Article 10). Additionally, a separate study using a

dataset of 768 patients developed seven ML algorithms and found that a neural network model with two hidden layers gave a 88.6% accuracy (Article 11).

Despite these advances, questions remain regarding the importance of individual symptoms and whether expanding the number of input features significantly improves predictive performance. This study aims to evaluate if a small set of hallmark symptoms is effective enough for diabetes screening, and whether the addition of secondary symptoms provides additional predictive benefit, using multiple machine learning models.

Materials and Methods

The dataset used in the study was found on Kaggle and is titled "Diabetes Risk Prediction". It contains 520 patient records with binary symptom-based clinical features as inputs and a binary outcome variable indicating the presence or absence of T2D. All input features are categorical and recorded as "Yes" or "No". The dataset is publicly available here: https://www.kaggle.com/datasets/rcratos/diabetes-risk-prediction.

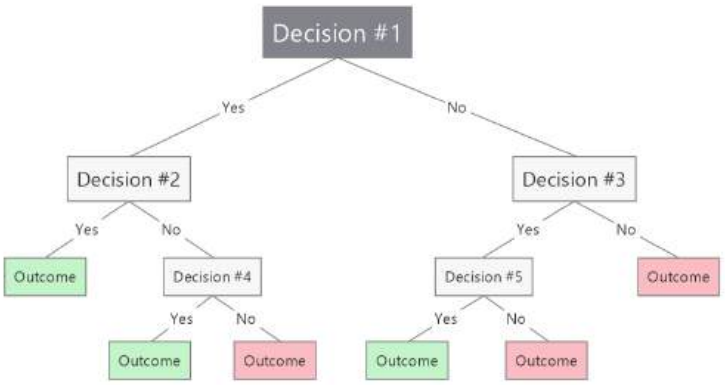
Two experiments were conducted to evaluate the effect of symptom selection on diabetes screening performance. Experiment 1 used three hallmark symptoms of diabetes: polyuria, polydipsia, and sudden weight loss. Experiment 2 expanded the feature set by adding two secondary symptoms: weakness and visual blurring.

All categorical inputs were converted from "Yes/No" to binary values (1/0). The outcome variable (class) was encoded as 1 for positive diabetes status and 0 for negative status. No missing values were present in the dataset.

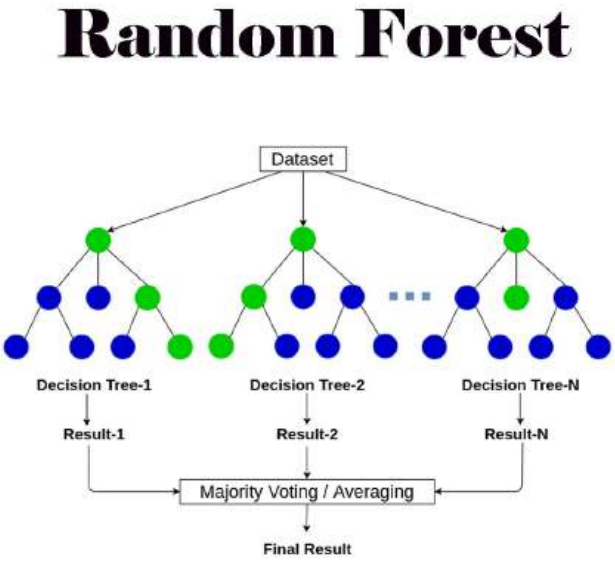
The dataset was randomly split into two parts: A training set (80%) and a testing set (20%). The dataset had a fixed random seed to allow for reproducibility between experiments. In the training set, the model was exposed to the data and built a system to give predictions. After the model used the training data, the testing set was then evaluated to determine the final performance of the model. To ensure a fair comparison between experiments, all model hyperparameters and the train-test split were held constant, and only the input features were changed between Experiment 1 and Experiment 2.

In this study, five different machine learning classifiers were trained and tested: a Decision Tree, Random Forest, Logistic Regression, Neural Network, and K-Nearest Neighbors (KNN). All of the models produced binary predictions (0 = negative, 1 = positive) for each patient record using the same train-test split and evaluation metrics.

The first model is the decision tree classifier. A decision tree classifier predicts outcomes by learning a sequence of yes/no rules that split the data into smaller and more uniform groups. The decision tree starts with the root node, which represents the initial decision and the variable that best separates the data at the start. At each split, the model chooses the symptom that best separates diabetic from non-diabetic cases. The final prediction is produced at a terminal "leaf" node. To reduce overfitting, the maximum tree depth was limited. The maximum depth of a decision tree is the maximum number of levels the tree can have, controlling how complex the tree can become.

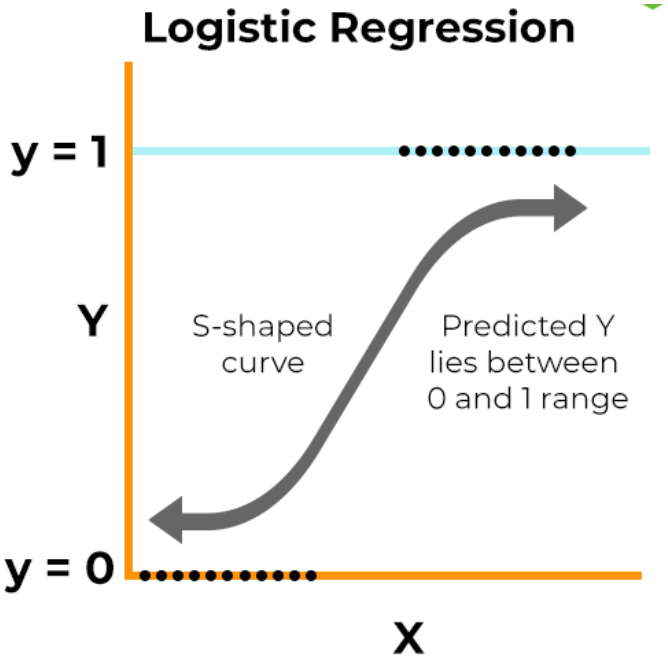


Another model used, similar to the decision tree, is a random forest model. The random forest classifier combines the predictions of many decision trees. Each tree is trained on a slightly different subset of the training data, improving the stability of the model compared to a single tree. For classification, the random forest model produces the final prediction by using majority voting across trees. The parameters of a random forest model include the maximum depths of the decision trees as well as the number of decision trees utilized. In this study, the number of trees and the maximum tree depth were held constant across experiments for fair comparison.

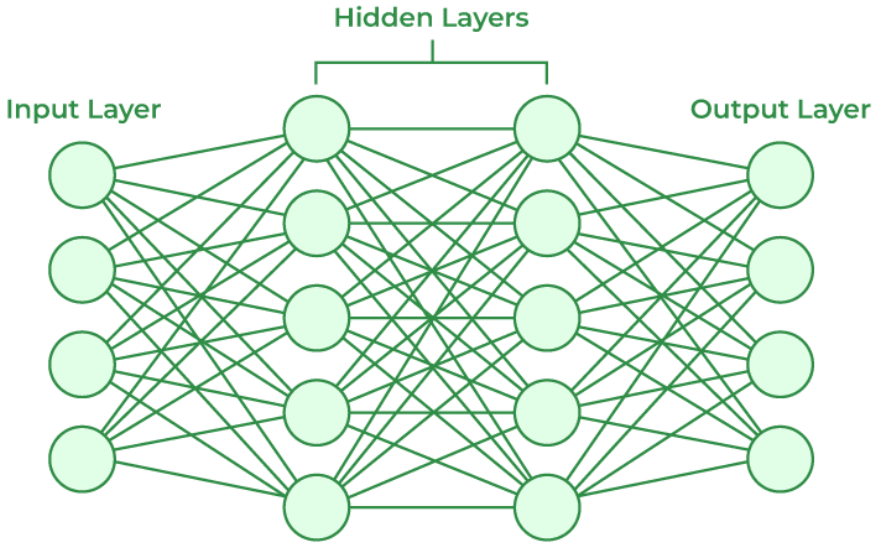


The next model is logistic regression. A logistic regression predicts the likelihood of diabetes by combining the input symptoms into a weighted sum, and then applying a logistic (sigmoid) function to convert the result into a probability from 0 to 1. The predicted class is

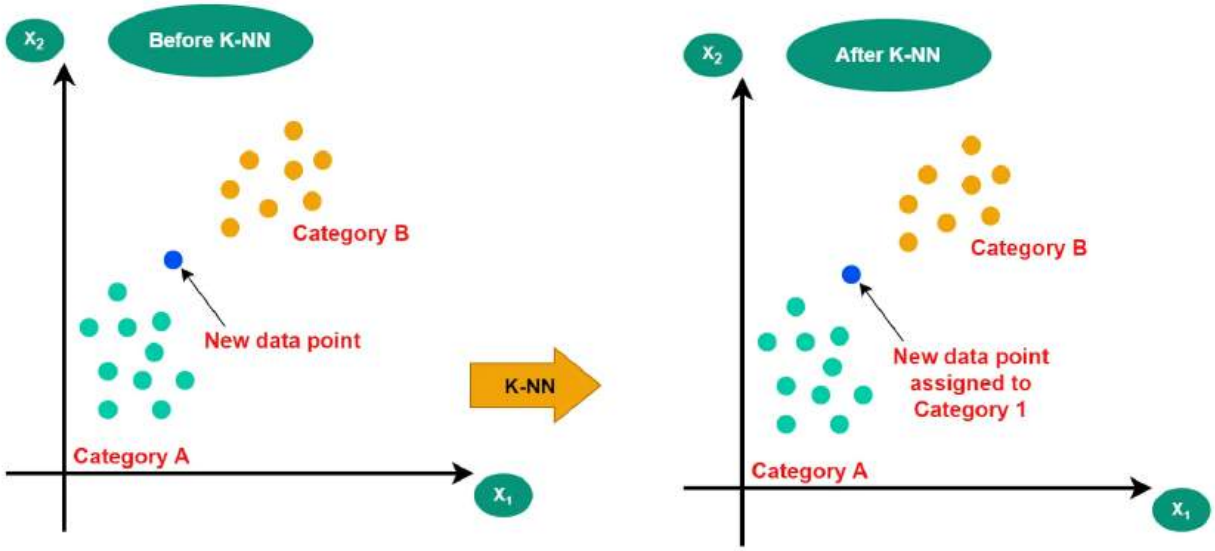
determined by a threshold (0.5). An important parameter in logistic regression is the number of iterations. The number of iterations determines how many times the model updates its weights during training to find the best solution. In this study, the number of training iterations was increased to ensure convergence.



The fourth model is the neural network model. A neural network is a model inspired by the way neurons communicate in the human brain and consists of layers of interconnected units called neurons. In the neural network, the input layer receives the symptom data, the hidden layers learn patterns by adjusting internal weights, and the output layer produces a final prediction. During training, the neural network repeatedly compares its predictions to the correct outcomes and adjusts its weights to reduce errors. The network was trained for a fixed number of epochs using a validation split to monitor generalization performance. Over time, the network learns complex relationships between symptoms that may not be obvious through simple rules.



Finally, the fifth model is the k-nearest neighbors (KNN) classifier. The model makes predictions by comparing a new data point to the most similar examples in the dataset. The model identifies the *k* closest data points (neighbors) based on similarity and predicts the class that appears most frequently among those neighbors. In this study, the KNN predicts diabetes status by examining patients with similar symptoms. If most nearby patients have diabetes, the model predicts a positive result. Unlike other models, KNN does not learn rules, but rather relies on stored examples to make predictions. The main parameter of the KNN model is the number of neighbors *(k)* considered when determining a prediction. The value of *k* was held constant across experiments to allow fair comparison of feature sets.



Each model's performance was evaluated using accuracy, precision, recall, and F1-score. Accuracy is the proportion of predictions that the model gets correct overall. For example, if the model makes 100 predictions and 85 of the predictions are correct, the accuracy is 85%. Precision measures how many of the cases predicted as positive were actually positive. When there is high precision, there are few false positives. On the other hand, recall measures how many actual cases of each class the model correctly identifies. Recall was considered particularly important in this study, as failing to identify individuals with diabetes could delay diagnosis and treatment. The F1-Score balances precision and recall into one value. It gives the overall quality of the classifier for each class when considering both false positives and false negatives.

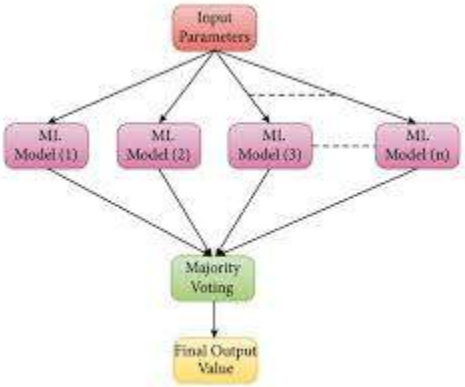
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

After training the different models (Decision Tree, Random Forest, Logistic Regression, Neural Network, and K-Nearest Neighbors (KNN), a Majority Voting Ensemble Method combines the predictions to produce a final decision. Each separate model produces a binary prediction for whether diabetes is present and the ensemble assigns the class that receives the most votes.



Results

The performance of the five machine learning models and a majority voting ensemble was evaluated across two experiments using different symptom sets. Experiment 1 used three hallmark symptoms of diabetes (polyuria, polydipsia, and sudden weight loss), while Experiment 2 added two secondary symptoms (weakness and visual blurring). The model performance was assessed on an independent test using accuracy, precision, recall, and F1-score. Results for each model are summarized in Tables 1 and 2, and a direct comparison of test accuracy between experiments is shown in Figure 1.

Table 1:
Model Performance Using Three Hallmark Symptoms (Experiment 1):

Model	Test Accuracy	Precision (Diabetes)	Recall (Diabetes)	F1-Score
Decision Tree	0.87	0.93	0.87	0.90
Random Forest	0.88	0.95	0.87	0.91
Logistic Regression	0.87	0.93	0.87	0.90
Neural Network	0.87	0.93	0.87	0.90
KNN	0.88	0.94	0.89	0.91
Ensemble (Voting)	**0.88**	**0.95**	**0.87**	**0.91**

In Experiment 1 (Table 1), all of the models had similar test accuracies, with a range between 0.87 and 0.88. Precision and recall for diabetes were consistently high as well, indicating effective identification of positive diabetes cases.

Table 2
Model Performance Using Expanded Symptom Set (Experiment 2):

Model	Test Accuracy	Precision (Diabetes)	Recall (Diabetes)	F1-Score
Decision Tree	0.88	0.95	0.87	0.91
Random Forest	0.88	0.95	0.87	0.91
Logistic Regression	0.87	0.93	0.87	0.90
Neural Network	0.88	0.92	0.92	0.92
KNN	0.88	0.98	0.85	0.91
Ensemble (Voting)	**0.88**	**0.95**	**0.87**	**0.91**

In Experiment 2 (Table 2), adding weakness and visual blurring resulted in minimal changes in overall performance. Test accuracy, precision, and recall remained relatively consistent with Experiment 1 across all the models. As there are small numerical differences between the experiments, no single model showed a substantial improvement with the expanded feature set.

Figure 1:

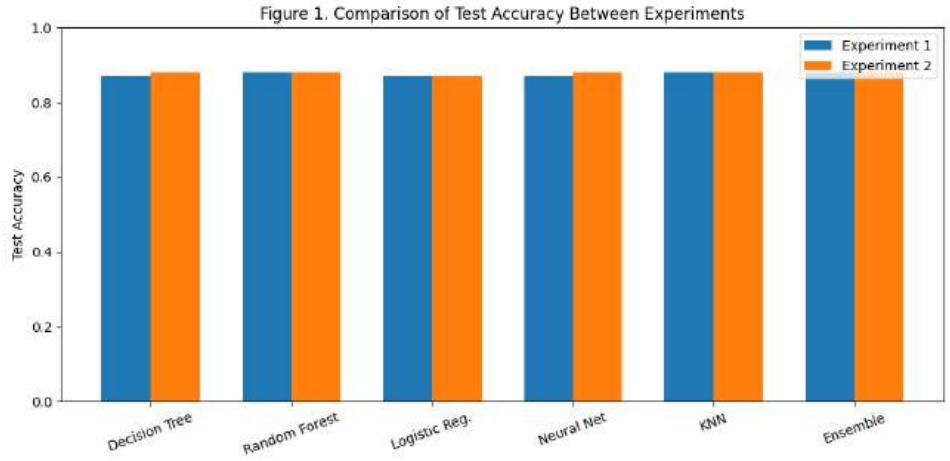


Figure 1 compares test accuracy across models between Experiment 1 and Experiment 2. The accuracy between the two experiments was minimal for all models, with overlapping performance across classifiers.

Discussion

The study evaluated whether expanding the number of symptom-based input features improves the performance of machine learning models for diabetes screening. Across the two experiments, five machine learning classifiers and a majority ensemble demonstrated consistently high performance when predicting diabetes status. Experiment 1 used three hallmark symptoms of diabetes- polyuria, polydipsia, and sudden weight loss- while Experiment 2 expanded the feature set by adding weakness and visual blurring. Although there was an increase in input features, the model performance metrics remained largely unchanged between the two experiments.

The consistently high performance seen in both experiments is likely due to the strong predictive symptoms. Polyuria and polydipsia are direct consequences of hyperglycemia, and sudden weight loss is due to the inability to use glucose and increased fat and muscle breakdown. These hallmark symptoms establish a set that distinguishes diabetic from non-diabetic patients, creating a strong underlying signal. Therefore, adding secondary symptoms in Experiment 2 did not provide substantial additional discriminatory power. As a result, test accuracy, precision, recall, and F1-score remained similar across both experiments.

Several limitations should be considered when interpreting these findings. First, the dataset was limited in size, leading to a lack of variability in performance and different models. Smaller datasets also increase the risk for overfitting, especially for more complex models. Additionally, the test set was small, meaning that small changes in predictions could have a larger effect on the metrics. Furthermore, due to the fact that the selected symptoms were already highly predictive of diabetes, the study design may limit the true potential of adding additional features.

The results raise several open questions for future research. How would the different models perform if the inputs were less direct symptoms or demographic and clinical variables such as age, sex, or obesity? Would training on a larger and more diverse dataset improve generalization and allow clearer differentiation between models? Moreover, how would performance be affected in populations with lower symptom severity or earlier stages of disease?

Future work could explore incorporating broader clinical and demographic features to assess whether expanded feature sets improve prediction in more heterogeneous populations. Models could also be optimized for screening applications by adjusting classification thresholds to prioritize recall and therefore minimize missed diabetes cases. Finally, validating the models on larger and independent datasets would help evaluate their generalizability and determine how well symptom-based screening approaches perform across diverse populations and real-world settings.

Works Cited

1. "Recent Advances in Diabetes Prediction Using Machine Learning." Journal of Medical Internet Research, 2024, https://pmc.ncbi.nlm.nih.gov/articles/PMC12392476/.

2. "Machine Learning Approaches for Diabetes Prediction." Journal of Medical Systems, 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC10591058/.

3. "Machine Learning Techniques for Diabetes Prediction." Procedia Computer Science, vol. 132, 2018, pp. 546–553, https://www.sciencedirect.com/science/article/pii/S1877050918308548.

4. "Prediction of Diabetes Using Data Mining Techniques." Journal of Diabetes & Metabolic Disorders, vol. 15, no. 1, 2016, https://www.sciencedirect.com/science/article/pii/S2001037016300733.

5. "Artificial Intelligence-Based Screening for Diabetes." PubMed, 2021, https://pubmed.ncbi.nlm.nih.gov/34902070/.

6. "Emerging Trends in Diabetes Research." PubMed, 2024, https://pubmed.ncbi.nlm.nih.gov/39168869/.

7. "Diagnosis and Classification of Diabetes Mellitus." Diabetes Care, vol. 37, suppl. 1, 2014, pp. S81–S90, https://diabetesjournals.org/care/article/37/Supplement_1/S81/37753/.

8. "Artificial Intelligence and Diabetes Prediction." Journal of Clinical Medicine, vol. 13, no. 3, 2024, p. 442, https://www.mdpi.com/1648-9144/60/3/442.

9. "Diabetes and Its Complications." The American Journal of Medicine, vol. 133, no. 7, 2020, pp. 818–826, https://www.amjmed.com/article/S0002-9343(20)30339-9/.

10. "Artificial Intelligence in Diabetes Diagnosis." Journal of Diabetes Research, 2020, https://pmc.ncbi.nlm.nih.gov/articles/PMC6942113/.

11. "Machine Learning Models for Diabetes Risk Prediction." Heliyon, vol. 7, no. 2, 2021, e06025, https://www.sciencedirect.com/science/article/pii/S2405959521000205.

12. GBD 2021 Diabetes Collaborators. "Global, Regional, and National Burden of Diabetes from 1990 to 2021, with Projections of Prevalence to 2050." The Lancet, vol. 402, no. 10397, 2023, pp. 203–234, https://doi.org/10.1016/S0140-6736(23)01301-6.

13. Kalyani, Rita R., et al. "Diagnosis and Treatment of Type 2 Diabetes in Adults: A Review." JAMA, vol. 334, no. 11, 2025, pp. 984–1002, https://doi.org/10.1001/jama.2025.5956.

14. "Diabetes Mellitus Overview." PubMed, 2022, https://pubmed.ncbi.nlm.nih.gov/36332637/.

15. Babel, R. A., and M. P. Dandekar. "A Review on Cellular and Molecular Mechanisms Linked to the Development of Diabetes Complications." Current Diabetes Reviews, vol. 17, no. 4, 2021, pp. 457–473, https://doi.org/10.2174/1573399816666201103143818.

16. Fang, M., et al. "Undiagnosed Diabetes in U.S. Adults: Prevalence and Trends." Diabetes Care, vol. 45, no. 9, 2022, pp. 1994–2002, https://doi.org/10.2337/dc22-0242.

17. Ogurtsova, K., et al. "IDF Diabetes Atlas: Global Estimates of Undiagnosed Diabetes in Adults for 2021." Diabetes Research and Clinical Practice, vol. 183, 2022, p. 109118, https://doi.org/10.1016/j.diabres.2021.109118.

18. Mugume, I. B., et al. “Performance of a Finnish Diabetes Risk Score in Detecting Undiagnosed Diabetes among Kenyans Aged 18–69 Years.” PLOS ONE, vol. 18, no. 4, 2023, e0276858, https://doi.org/10.1371/journal.pone.0276858.

19. Ruige, J. B., et al. “Performance of an NIDDM Screening Questionnaire Based on Symptoms and Risk Factors.” Diabetes Care, vol. 20, no. 4, 1997, pp. 491–496, https://doi.org/10.2337/diacare.20.4.491.

20. Aglarci, A. V., and F. Karakurt. “Symptoms Affecting the Development of Diabetes: Analysis of Risk Factors with Data Mining.” BMC Medical Informatics and Decision Making, vol. 25, no. 1, 2025, p. 319, https://doi.org/10.1186/s12911-025-03159-5.

21. U.S. Preventive Services Task Force. “Screening for Prediabetes and Type 2 Diabetes: U.S. Preventive Services Task Force Recommendation Statement.” JAMA, vol. 326, no. 8, 2021, pp. 736–743, https://doi.org/10.1001/jama.2021.12531.

22. American Diabetes Association Professional Practice Committee. “Diagnosis and Classification of Diabetes: Standards of Care in Diabetes—2026.” Diabetes Care, vol. 49, suppl. 1, 2026, pp. S27–S49, https://doi.org/10.2337/dc26-S002.

23. World Health Organization. “Noncommunicable Diseases.” World Health Organization, https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases. Accessed 26 Jan. 2026.

From Elegy to Tragedy: The Enduring Poetics of Exile from Ovid's Tristia to Racine's Phèdre By Elle Lee

> If only death had pre-empted me before my exile,
> or my punishment had been deferred
> to the hour of my death! Till lately I could have concluded
> my days without taint; but now I've been granted
> life -
> to die an exile (Tristia book III part III lines 33-3)

Abstract

Exile has long been a defining literary motif, often stemming from personal experience and evolving from physical and political punishment to emotional and intangible portrayals. Ovid's Tristia transforms banishment into a poetic identity in which language becomes both a haven and a deepening wound. This paper traces the evolution of exile's representation from Ovid's elegiac poetics to Racine's 17th-century tragedy Phèdre, focusing on how each writer renders themes of shame, death, and communication. Racine inherits Ovid's emotional poetics of exile but redirects it inward, transforming exile into a psychological and metaphysical ordeal. In Racine's dramatization, passion leads to shame and eventual self-banishment. Through close readings of diction, imagery, and form across Tristia and Phèdre, I argue that exile functions as a literary device that shapes both narrative structure and the emotional texture of language. In Tristia, Ovid presents expulsion to be a destructive and divisive event, leaving writing as his final means of survival and comfort. In Phèdre, Racine inherits this poetics of exile but adapts it to drama: now shifting from written catharsis to verbal dialogue in order to fit the structure of a play, which externalizes Phèdre's internal displacement. By juxtaposing Ovid's solitary elegy with Racine's dialogic stage, this paper reveals exile's unifying and enduring capacity to express the human struggle for belonging. Ultimately, the comparison underscores exile's continual transformation across genre, language, and time, revealing its enduring resonance in modern narratives of displacement and estrangement.

Introduction

Exile has long been a universal theme throughout literature. In ancient eras, the concept of exile was a more formal and concrete banishment. With time, exile has complicated and gained dimensions, providing nuanced modern examples such as immigration and deportation. As exile has transformed, so has its representation in literature. Yet, the complex emotional toll that banishment takes on individuals has persevered. Exile often traps its victims in such suffocating solitude that, as Ovid noted, they wish that death "had pre-empted…" them "before [their] exile." Exile comes with intense emotional baggage consisting of guilt, self-hatred, loneliness, and feelings of betrayal, often leading to this romanticization of an unmarred death. Language frequently serves as their asylum during this tortuous time. Narratives written to serve as refuges during estrangement have traversed time and domains to unite people of endlessly varying backgrounds. In his tragedy *Phèdre*, Jean Racine uses a dramatic format to build upon

Ovid's poetic conception of physical exile in *Tristia* and present an internal, emotionally-centered portrayal of exile.

Ovid's *Tristia* serves as a major point of reference in the evolution of exile as a theme and its representation in literature. In AD 8, renowned poet Publius Ovidius Naso was exiled from Rome to Tomis on the Black Sea for an unclear offense against the Emperor Augustus, which was in some part a result of his scandalous poetry. Having been punished for his art and stranded in a civilization that turned language into an isolating factor, Ovid still turned to writing as a haven. He channeled his turbulent experiences into *Tristia*, a book of autobiographical poetry containing the catharsis of his exile (Green, xxxix). His work serves as an influence upon exilic literature extending beyond the barriers of time, geography, and even genre, continuing to impact countless literary pieces.

Ovid's interpretation of exile manifested in countless lives beyond his own, extending to seventeenth-century France. In 1661, French playwright Jean Racine estranged himself from his community of Jansenist friends and mentors in order to pursue his passion for theatre. His work in the theatre was frowned upon by the Jansenists, so his choice to write as a dramatist was a finalizing one. His celebrated play *Phèdre*, first performed and published in 1677, echoes the same themes of isolation that he confronted in his experience of exile (Tobin, "Jean Racine"). His character Phèdre grapples with the shame for and regret of her betrayal of her family, and her journey and eventual suicide represent a voluntary exile. Even through these distinct elements, Ovid's exilic influence remains prominent in Racine's work. While Ovid writes exile as an imposed condition, Racine turns it into an internal state of guilt and desire.

The theme of exile is one that has transcended time, evolving primarily from Ovid's *Tristia*. Many more famed pieces of literature reflect this trajectory, such as *The Odyssey* by Homer, but *Phèdre* is unique in its ability to demonstrate these echoes of Ovid's influence through a different language, genre, and cultural setting. Racine expands Ovid's conception of exile by turning it inwards into a psychological and moral punishment, which is represented differently through their unique form choices. Ovid's work of autobiographical poetry offers obvious authenticity, and while Racine's play is fictional, his manipulation of dialogue serves the same purpose of exposing vulnerability. Therefore, this combination of compositions is essential in investigating the true impact of Ovid's work on later themes of exile, which is present through complications explored over time.

There are clear differences between Ovid's representation of exile and that of Racine, Ovid's connections to themes of loss and language reverberate in *Phèdre*. However, the way that they manifest alludes to expected shifts that occurred between *Tristia*'s and *Phèdre'*s conceptions. Ovid's exilic work serves as a baseline from which *Phèdre* and other works could be composed. Exploring the contrast and development of concepts from Ovid's *Tristia* to Racine's *Phèdre* presents a unique model for investigating further pairings of literature. Therefore, comparing the nuance in their representations offers a meaningful understanding of this evolution that can occur from a shared theme, despite the passage of time, varied genres, and diverse language.

Ovid set the stage for endless interpretations of exile, beginning with his more simple and nonfictional poetry concerning his experience in extreme physical banishment as punishment. Exile has been conceived in an unending number of ways over time, manifesting in the modern world in occasionally more voluntary forms such as migration and immigration, but in Ovid's case, exile confined him permanently to Tomis on the Black Sea. Though Emperor Augustus was a patron of the arts, he did not hesitate to discipline Ovid. His influence created a suffocating exile for Ovid through not only cutting Ovid off from his home but also destroying many of Ovid's relationships through fear of disrespecting the emperor (Augustus). Ovid's banishment was therefore one characterized by a sense of total isolation and contempt, a jarring contrast to his previous public favor. In exile, Ovid faced a disheartening transition period in which he had to grieve his past life and accept his new reality. He mulled over his situation through endless compilations of poems, writing:

> While I walked safe still, I yearned for recognition,
> was on fire to make myself a name;
> but now, let it suffice me not to detest the poems,
> the pursuit that undid me: it was my own wit
> brought me to exile (Tristia book I Lines 53-57)

Ovid's Poetics of Exile

Ovid's reflection back to his past, during which he "walked safe still," demonstrates the intense nostalgia that came with his exile. In his ostracized state, Ovid naturally yearned for and romanticized his memories of living in his community. He grieved the luxury of feeling as if he was "on fire" with passion, but he did not neglect his culpability in his situation. Ovid recognized his transgression and admitted guilt in describing his writing as having "undid" him. His choice in describing his work as having gone so far as having "undid" him hints to the increasing issues Ovid had with understanding himself outside of, overshadowed by, and in conjunction with his identity as a poet. With his explicit shift in tense, Ovid acknowledged the divide that he believed exile imposed on his life, dictating a time of before and after disaster. Yet, he persevered in declaring that "now," he would "not…detest the poems" because he knew that writing was not something he could stand to abandon in resentment when faced with such loneliness. He found writing in exile to be both a refuge and a further agitator of his loneliness. He used his poetry as his only connection back to Rome, and it served as a safe space of true vulnerability in which he felt understood. Yet, he eventually became disillusioned with his writing as its tenuous link back to his past community became not a solace, but a scar. His careful mechanical choices in his poetry created the vulnerable story of a poet left abandoned by his readers, which is reflected in reverse in Racine's *Phèdre* through Phèdre's eventual abandonment of her community and audience.

Ovid's work presents the eventual destructive effects of exile on one's psyche. In the early days of Ovid's banishment, he maintained exceptional faith and perseverance. He was

adamant in the injustice of his predicament, and though he conceded that he did have some fault in the situation, he was relentless in his appeal for a better exile. Ovid did not let himself become discouraged by his loneliness, and in his second book, declared to Augustus that "you yourself forbid hope, yet I'll hope for ever -/this one thing I can do against your will (Book II, Part 1, lines 145-146). His determination and "hope" were his ways of combatting the oppression he faced, and therefore, he did not lose his belief in himself and his writing in exile. Yet, as time went on, his dehumanizing isolation eroded these seemingly indestructible confidences in himself. Worn down by more and more merciless seasons of exile, Ovid ultimately succumbed to the relentless feelings of shame and self-hatred that arise in banishment. Life in solitude left to contemplate the cause of his exile and grieve his past life alone demoralized Ovid, who finally admitted these sentiments in his fifth installation of *Tristia*, in musing "… How I wish/that my *Art*, which destroyed its unsuspecting maker,/had been turned to ashes first" (Book V, 12, Lines 66-68). Throughout *Tristia*, Ovid was very adamant about his belief that the content of his poetry was not a just reason for his exile. The fact that at the end of Book V, he went so far as to wish his writing "had been turned to ashes first" emphasizes how successfully exile had broken his seemingly impregnable spirits and dedication to his art. Ovid's choice in wishing his poetry had been "turned to ashes first" rather than simply never written or unreleased underscores how absolute he desired his past work to be destroyed to have saved him the torture of exile. This sense of hindsight regret for not having demolished passion is, too, reflected in *Phèdre*'s laments. Here, Ovid no longer defended his work, and instead he blamed it for having "destroyed its unsuspecting maker," still attempting to absolve himself from culpability through separating himself from his writing. He spoke about his poetry not only as an entity separate from himself, but in a detached tone that even put distance between him and the version of himself that was originally hurt by this situation. This estrangement from such a vital element of his being mimics the physical one that Ovid experienced on Tomis.

Not only did exile lead Ovid to begin to resent his writing, but in time, he started to despise himself. Ovid, once the celebrated and distinguished poet, now reduced to disgrace, could only maintain his self-worth for so long. His pride diminished, and by the time he composed his fifth book, he defeatedly remarked that there was "No hope for me either: I wasn't that much to start with,/but I can never return/even to what I was. …" (Book V, 12, Lines 29-31). Ovid provided a sharp contrast from his declaration of "hope for ever" (Book II, Part 1, line 145) in his first edition of *Tristia*, finally accepting that there was "No hope" for him. Ovid commanded such raw repetition of these motifs to authentically depict the harrowing despondency that can hollow someone in exile. Ovid went even further than pessimistically positing upon his future and even condemned his past self as not having been "that much to start with." His choice in describing himself as unexceptional in his lack of being a simple "much" demonstrates his deep contempt towards his past self, a symptom of his exile. This intense self-loathing retained echoes even in Racine's work, which he chose to emphasize in accordance with his dramatic structure. Ovid's deliberate structure, too, plays a role in expressing the great impact that this time of ostracization has had on his respect for himself as a poet. Ovid's choice

in adding a line break before “even to what was” emphasizes the statement in how he felt that though he had sunken lower than ever before, reverting to the version of himself that he used to be would not even be that much of an improvement.

Ovid’s expression of this degradation of both his talent and sense of self offer powerful examples of conveying the long-term consequences of exile through intentional and authentic structure. Following these motifs in his work is just one example of a path he forged, but the entirety of his exilic work sets up this inspirational foundation for future exilic works to be built off of. Ovid provided a strong framework to reference in emotionally communicating the raw turmoil of exile. Ovid crafted a vivid look into how his calling to poetry is impacted by isolation.

Ovid went on to write autobiographical poetry, yet his story was not centered around his personal struggle, but rather his identity as a poet in exile. Critic B.R. Fredericks encapsulates this concept in offering that “Ovid's focus is on his career as a poet, not on his life as a man” (Fredericks 143). In isolation, Ovid became so reliant on writing as an outlet that he became consumed by grappling with the identities of the “great poet” that he once occupied and “the exile” that he was now absorbed into. Ovid continually referred to himself as an “exile” and constantly brought up how famous and adored his writing once was. He obsessed over his identity as a poet rather than a struggling man in isolation because the version of himself that he most associated with his home in Rome was the poet. Therefore, he writes with a focus on how this banishment impacts his experiences as a poet to feel more connected to his past self in Rome. Ovid grew to accept the continued solitude that he will receive in writing, contemplating:

> […] though my Muse (a guest despite misfortune)
> still contrives to return to verse, to her old rites,
> there’s no one to whom I can recite my poems, no one
> whose ears are attuned to Latin verse,
> so I write, and read, for myself (there’s no other option) -
> am my own judge: my work is safe with me. (Book IV, Part 1 Lines 87-92)

Ovid expressed that though he yearned to “return to verse,” he was left with “no one” to appreciate it. In his exile, not only was he cut off from the people he loved, but from any collection of people who could even speak the same language as him and form a basic understanding of his art. Ovid admitted that his continuation of his work was not even just because he chose to focus on his self-improvement to spite the people that cast him away to isolation, but because he believed there to be “no other option.” Especially as someone whose life was built off of acclaim from writing, the harsh removal of Ovid’s audience was quite the blow. Yet, he persevered in calling his writing “a guest” despite the ramifications it had on his life, highlighting his defense of his creativity for his own sake. His writing provided him with a sense of comfort through familiarity, his mastery of Latin reminding him of his past success and relationships in Rome. Ovid described how his “Muse” was still captivated by “her old rites,” suggesting this connection between Ovid’s writing and the way that it measures and marks time

in his life that he could not sever. Ovid was also aware that because of his isolation, he was free to be his "own judge," and with nobody left to criticize or censor his work for being too scandalous, his work was "safe with" him. He found solace in this silent response to his art, having already lost anything that could be taken from him in punishment for his self-expression. Even though Ovid recognized that his writing was a major factor in what caused his sentence of banishment, he remained determined not to let that experience discourage him from continuing to write through his exile.

Despite the refuge Ovid found in his manipulation of language, he found it to be a further ostracizing factor not only from the people in his home, but from the people of Tomis. In Tomis, the locals spoke a different language, so his feelings of displacement were amplified. Critic Sabine Grebe clarifies that "The poet experienced Tomis as being situated beyond this known world in an unknown and unknowable environment" (Grebe 496). Grebe offers insight into how Ovid considered Tomis to be full of the "unknown" in terms of culture and terrain, but more significantly, "unknowable" because he perceived the language barrier to be an insurmountable hurdle. Suddenly left in this new community, Ovid did not even believe initially in the possibility of growing a relationship or at the very least linguistic understanding with the people of Tomis. He immediately decided that communication with the peasants would be a futile pursuit and was left without anyone capable of conversing with him or enjoying his work near him. Coming from a past characterized by such reverence for his ability to reach people through words, Ovid found this new reality intolerable. Frustrated, he wrote:

> Here, *I'm* the barbarian, understood by no one,
> and these stupid peasants mock my Latin speech,
> slander me to my face with impunity, on occasion
> (I suspect) laugh at my exile… (Tristia Book V Part 5 Lines 37-40)

Ovid's use of italics in describing himself emphasizes just how jarring he found life in Tomis to be, stripped from his previously adored ability to communicate. His harsh language in recounting the people's perception of him, including "slander," "impunity," and "barbarian," demonstrate how deeply ostracized Ovid felt. His poetry became the only setting in which he could express sensitivity, so the fact that he was so repetitive in his explanation of his idea of the locals' perception of him displays how acutely affected he was. He refers to the peasants as "stupid," suggesting that in his separation from those who understand his language, Ovid was so misunderstood that even "stupid" people are able to mock him, he who was once regarded as a literary genius. In the ambiguity and confusion that this lingual disconnect left him in, Ovid, alone with only the echoes of his negativity filling his mind, immediately interpreted that the people of Tomis must have "laugh[ed]" at his exile. This deep severance from an understanding audience leaves Ovid spiraling through his pen, writing in his native language to cope with his pains of loneliness and guilt. Ovid's complicated relationship with language is a key feature in his exilic work that outlives him.

Ovid's work describing exile has served as a time-enduring guide in examining the contours of exile. Ovid's trajectory offers an authentic view of the intricacy of ostracization and the messiness of the victim's emotional state. Ovid provides a framework off of which infinite exilic stories can be written, with unique details. He accurately crafts a story full of realistic contradictions to convey the fickle perspective of someone traumatized by exile before their spirits eventually broke. Ovid's representation considers the nuance of isolation that manifests in questioning loyalties, clinging to any forms of authentic communication, and even considering death to be a salvation. These same patterns appear in exilic poetics from entirely different backgrounds, being expressed with varying interpretations but maintaining the same basic elements considered in the foundation Ovid built. Ovid's exploration of relationships damaged by exile's shame provides a foundation for comparing his interpretation with Racine's more emotionally-centered Phèdre.

Racine's Tragic Exile and Inner Banishment

If Ovid turned exile into a poetic identity, Racine reimagines it as a psychological ordeal culminating in self-banishment. Racine's work in *Phèdre* reflects the sort of self-inspired exilic themes as Ovid's *Tristia*. Racine's experience distancing himself from and subsequently being held in contempt by his Jansenist community offers insight into the authenticity of the exilic themes in his work. As a playwright, Racine's work was naturally more dramatized and less abstract than Ovid's poetry. Ovid's elegiac poetry is imbued by an authentic mourning of his past and potential, and therefore takes on a more serious tone. Yet, because of Racine's choice to write a tragedy, he had the obligation to amplify the emotions and expressions of his characters in order to make them more understandable for the performers and audience in addition to readers. Phèdre must speak of her shame publicly, whereas Ovid's physically isolated lament is private, which offers different experiences for the audience to understand the depth of each character's exile. However, even through a different form, Racine expresses the same complications of loneliness and grieving loyalty as Ovid through an immaterial landscape. While *Tristia* centers on physical exile and its emotional effects, *Phèdre* depicts a psychological banishment that leads to physical self-destruction.

Phèdre's exile is primarily an internal one out of shame for her forbidden passion. Contrary to Ovid's very physical experience with displacement, Racine takes Phèdre's dramatic tale as an opportunity to explore a more evolved and complex mental banishment. As epitomized by Critic Ronald W. Tobin, "Geographical displacement is, in Racine, often converted into spiritual displacement: captives, strangers, exiles, wanderers in one sense or in all senses are the characters of his plays" (Tobin 2). In Racine's work, Phèdre grapples with her traitorous love for her step-son, Hippolyte. As she comes to terms with these disgraceful desires, she endures a more isolated journey in this "spiritual displacement" that Tobin describes, taking on each of these roles of "captive" of her emotions, "wanderer" in a territory of indecision, and eventual "exile" in death. On this impossible path, Phèdre becomes consumed by this self-repulsion, and in her conversations with confidants, admits:

J’ai conçu pour mon crime une juste terreur;
J’ai pris la vie en haine, et ma flamme en horreur.
Je voulais en mourant prendre soin de ma gloire,
Et dérober au jour une flamme si noire:” (Racine 52)

Now I am seized with terror for my crime;
I hate my life; my love is horrible.
I wish, in dying, to preserve my fame,
And hide from light of day a love so black.” (Rawlings 53)

Phèdre demonstrates her shame for her passion in her description of having to face a resulting “juste terreur,” describing the fear as deserved. Rawlings translates Phèdre’s line as “I hate my life; my love is horrible,” but a closer rendering—“I took my life in hatred, and my flame in horror” better captures her active self-condemnation. Here, Phèdre acknowledges her own agency in creating a reality defined by “haine” and “horreur.” This dramatic arc exhibits Racine’s work in dramatizing exile as a psychological process by identifying its advancement through emotional marks. Phèdre highlights these points by injecting her words with chagrin in her portrayal of her feelings as “black.” In Racine’s original work, Phèdre depicts her “love” as “une flamme,” a flame, which conjures images of a powerful and all-consuming force of destruction if not extinguished early enough. She darkens the image by calling the flame “black” and struggles to “hide” it from the “light of day,” a futile effort since both evoke brightness and exposure. These repeated motifs of visibility point to Phèdre’s desire to avoid the drama that would arise should the people around her and even the public learn of her deceit. Especially as the wife of the king, Phèdre must consider the impact of her secret being exposed not only to the people immediately of effect to her, but to everyone in her larger community. Therefore, having to face the possibility of all of this judgement along with the weight of her personal guilt, Racine offers an accurate representation of Phèdre’s inner struggle through these dark metaphors and references.

These intense visions reflect Phèdre’s internal spiral of shame as she navigates her unfortunate situation, which drives her to eventually seek death as an exile. Racine structures Phèdre’s reflections as a logical progression toward exile. Constrained by the form of drama, he nevertheless conveys her complex internal struggle through concise, nuanced dialogue with trusted confidants. Through his use of dialogue, Racine lets Phèdre’s ramblings provide a clear look into the true pattern of activity in her mind. Moving down along the stanzas, Phèdre’s focuses shift from “terror” to “hate,” and at the halfway point two lines down, she begins to reframe her thinking to prioritize finding a solution. She remembers to consider her personal wants, and as she recalls how much her reputation means to her, she reaches the resolution that death will be her solace. She refers to wanting to “preserve” her reputation in death, similar to Ovid’s yearning for an unmarred death. Trapped in this inflamed situation of which she feels that

she has lost control, Phèdre expresses thoughts of death as an opportunity to take back control over her story. Phèdre's view of death begins as one romanticizing the escape she would have if she were to die with her legacy preserved, but when the possibility of a shameless death becomes compromised, she shifts to seeking death as a form of self-banishment. Racine's work in describing Phèdre's journey in her view of death as banishment offers more opportunities to compare nuanced and more obvious shifts from Ovid's depiction of the decline of hope in exile.

Similar to how Ovid suffered the tensions and eventual ruptures of relationships following his exile, Phèdre's dark passion severs many bonds and leads her to spiral further towards thoughts of death. When Phèdre expresses her traitorous passion, she faces not only her own potent self-hatred, but the disapproval of others that manifest in the test of their loyalty to her. As Racine conveys Phèdre's story as a play, her turmoil is shared through vulnerable moments in dialogue, always directed towards another. Therefore, she directly confronts the feelings of treason she feels are mirrored back at her when her loved ones respond to her offense. She eventually demands:

> Ou si d'un sang trop vil ta main serait trempée,
> Au défaut de ton bras prête-moi ton épée.
> Donne. (Racine 86)
>
> Or if your hatred envy me a blow
> Of such sweet torture, if blood too vile
> You think would therefore drench your hand, then give,
> Give me, if not your arm, at least your sword!
> Give. (Rawlings 87)

When Phèdre feels a lapse in loyalty in someone's refusal to offer their "arm" in this time of need, she is driven to thoughts of death to the point of instead asking for their "sword." She refers to the disloyalty of her friends with a tone of disdain and resentment, sarcastically stating that she understands their stance of abandonment out of fear of "drench[ing their]... hand" in the "vile" blood she has soaked herself in. The fact that Racine writes Phèdre to demand the instrument of death not only once, but twice, in separate stanzas, emphasizes how intent she is on pursuing death. Racine's choice in having Phèdre describe the culpability by association that her friends would have to face as being a force so powerful that it would "drench" them additionally depicts her extreme shame.

Phèdre's case comes with more baggage of self-hatred than Ovid's, as her situation was entirely her fault and left without ambiguity, whereas Ovid's was more externally rooted and provided him with a figure other than himself to blame and resent. Therefore, her reckonings with death grow to be a final act of self-destruction resulting from her self-repulsion. When Phèdre finally confronts the death she has been preparing herself for, she exclaims:

Hélas! du crime affreux dont la honte me suit
Jamais mon triste cœur n'a recueilli le fruit.
Jusqu'au dernier soupir de malheurs poursuivie
Je rends dans les tourments une pénible vie. (Racine 128)

Alas, my sad heart never plucked the fruit
Of pleasure from the frightful crime of which
I stand accused and dogged by shame. Alas,
To my last gasping breath by griefs pursued,
I here surrender my tormented life. (Rawlings 139)

Phèdre emphasizes how her shame "[se] suit," or follows her, despite everything she has done to try to rid herself of it. She recognizes how she has been stuck in such a hopeless cycle in admitting that she was never rewarded the "fruit of pleasure" that she secretly hoped to achieve despite the scandal of her love. Finally having reached acceptance, she resolves to self-exile and "surrender" herself. The despair embedded in her description of this sort of banishment as "surrender" further stresses how this choice was driven by intense self-hatred.

Although Phèdre's situation differs from Ovid's in that her conception of exile was a more self-perpetuated and individual decision, Racine maintains the same literary connection between the exiled and community as Ovid. Ovid used his autobiographical poetry to connect himself to the people he cared for, however one-sided it was. Yet, because Racine's mode in telling Phèdre's fictional story is so technically different from Ovid's poetics, relying deeply on dialogue to fill that need for authenticity and subsequently forcing Phèdre's character to traverse her relationships following this confession. Her primary confidant in these endless moments of torment is Œnone, Phèdre's beloved nurse. Norman Levin emphasizes that "the role of Oenone, in the absence of whose misconceived and less-than-divine intervention the heroine would never have reached her divinely appointed doom, ought not to be left out of account" (Levin 61). Through their conversations, Racine creates the same forum for open exorcism of traumas and chaos for Phèdre that Ovid found in his poetry. Her inner thought processes become clear to Racine's audience as she is influenced by Œnone's advice and verbalizes her responses. Secondary characters in *Phèdre* such as Œnone function as mirrors for Phèdre to interact with; Conversations with others serve as moments in which Phèdre can explicitly self-reflect and reveal her true feelings to Racine's audience. Her relationship with Œnone specifically emphasizes Racine's reliance on Phèdre's community in expressing her journey of exile.

Œnone's potent loyalty and eventual tandem death by suicide highlight Racine's divergence from Ovid's complete isolation in exile. Racine's and Ovid's differing choices in structure contrast further when examining the narrative voices of each tale. Ovid's sole voice in his writing underscores his true abandonment in Tomis, whereas Racine's play-structure demands conversations between each character, offering further insight that in turn makes the reader more conscious of the community Phèdre is surrounded by. Racine's use of confidants

creates a shared, communal dimension to exile that is a sharp contrast to Ovid's solitude. Despite the added stress of questioning what that community might think of Phèdre, it remains there, providing her with a sense of normalcy as she faces exile in addition to the extreme loyalty of her select few. Racine provides an alternate facet to displacement, exemplifying the complications that the concept cultivated over time, depicting an experience softened by the true loyalty of a friend willing to follow an exile to their doomed sentence. Although Ovid occasionally wrote in appreciation of his few friends remaining loyal to him back home and his heartbroken wife, his gratefulness was overshadowed by his grief of relationships he had previously believed to be sound, and he had no one willing to brave exile just to be by his side. Ovid's bitter reception of this perceived betrayal provided an opportunity for him to express his frustrations. In his speech about these relationships, these people became vessels for Ovid to externalize his internal shame and sense of abandonment in a way of confronting it without the entire weight being on him. Ovid felt safe revealing this projection of his feelings onto others because of how isolated he and his writing was, as he could not receive responses to his work. Therefore, his literary interpretation of exile represents a more desperate call for understanding and comfort that Phèdre seeks in her conversations with Œnone, but she receives a response due to the mechanics of Racine's playwriting along with the physical distance between Ovid and his fading relationships.

Racine follows Ovid's basic arc of exile of a person being persecuted for their passion, initially turning to that passion for comfort in their distress, and eventually resenting that passion and themself. In Ovid's case, after being banished for his writing, he found comfort in his poetry, before eventually expressing that he felt that it had ruined him. Phèdre faces torment for her incestuous desire for Hippolyte, can no longer stand it and admits her feelings to him, and when he rejects her she banishes him and eventually banishes herself in the eternal exile of death. In this sense of ostracization, both grapple with grief of their simpler pasts, shame, and eventual self-hatred. In their resentment of their passion and of themselves, Ovid and Phèdre share longings for a death before their offenses had complicated their lives. In considering this alternate reality, Ovid mused "And - last but not least - I'd have died, as I lived, without scandal:/now my life is shamed by its very ordeal" (Tristia Book IV Part III Lines 47-49). Ovid's description of his entire "life" being "shamed" by his transgression emphasizes the sense of irreversible and all-encompassing guilt that Ovid and Phèdre experience. Yet, Phèdre is so consumed by this idea that her life is ruined by this mistake that she takes her wish for death past nostalgia and into the future. As Phèdre is facing the turmoil of her internal sense of self-condemnation, she cannot bear to take the additional blame from the community that she is surrounded by. Therefore, Phèdre concludes that she must exile herself in death, declaring:

> Mourons. De tant d'horreurs qu'un trépas me délivre.
> Est-ce un malheur si grand que de cesser de vivre?
> La mort aux malheureux ne cause point d'effroi.
> Je ne crains que le nom que je laisse après moi. (Racine 102)

I am resolved to die, for death alone
From all this horror can deliver me
Is it so great calamity to die? Death is not feared by those in misery.
I only fear the name I leave behind." (Racine 103)

Comparative Analysis: From Elegy to Tragedy

Phèdre cites the same "misery" that Ovid conveys in his dejected tone, which, in her case, is enough to drive her to choose "death." The legacy of "the name…[she] leave[s] behind" is the summary of "life" that Ovid feared he destroyed. The tonal nuance between these ideas result from the genre dissimilarity. Ovid's elegaic and autobiographical rhythm conveys this despair in a more casual manner, using em dashes to indicate his more informal communication. Racine, on the other hand, deliberately paces his tragic cadence to have dramatic stanza and line breaks, which in turn inflates the emotion of the passage. Ovid's motifs of death are sustained in Racine's work, but Racine expands Phèdre's relationship with death in his dramatization of her exile.

Both Ovid and Phèdre endure the toxic push and pull with their passions, but it manifests differently between Ovid and Racine due to their contrasting genres. Ovid's passion is his writing, which produces a more clear and physical representation, while Phèdre's passion is for Hippolyte, and therefore its representations are more intangible. Phèdre's exile mirrors the emotional nature of her passion, as it becomes a sentimental turmoil, which is because Racine writes in a dramatic style to follow his theatrical form. Racine's writing as a drama logically requires dialogue, so Phedre must express the same cloud of feelings that Ovid explored in his writing through her conversations with confidants to convey her thoughts to Racine's audience. While Ovid's mental release is sent overseas and without reply, Phèdre's articulation of her experiences is received not only by the people around her, but by the audience observing her performance. Phèdre's community in contrast to Ovid's true isolation in navigating his exile is an impact of Racine's deviating form choice. Racine follows Ovid's fundamental themes of exile, but as was popular during his time, Racine twisted it to produce a classical tragedy, resulting in a modernized depiction of exile.

Conclusion

Ovid's seminal emotional themes of exile in *Tristia* reverberate in Racine's *Phèdre*, though Racine's adaptation wields a dramatized mode of writing to explore a more complicated and metaphysical model of banishment. Due to the technical differences in Ovid's and Racine's choices in presenting their works, the audience reads the elegiac *Tristia* as an honest and isolated view of total physical exile, while *Phèdre* is presented as a dramatized demonstration of exile that therefore requires conversation with the community and creates a more immersive experience for the audience. Comparing *Tristia* and *Phèdre* reveals how literature across time and language revisits exile with new nuance. Both works advance a common theme that links emotional displacement to creative survival. Exile's literary power is in how form becomes a site

of survival for its protagonists, which is shown through its durability as both a lingual and emotional condition. Banishment has always been and will always be thematically relevant in society through its endless transformations. Displacement in the modern world is rendered in physical matters of deportation and immigration, and takes on the metaphysical in virtual ostracization through social media. Yet, the constant among these diverse experiences is the expression of survivor's story, which transcends all disjunction across periodical, territorial, or communicative circumstances. Further research may find success in scrutinizing current applications of exile and their heavily online articulations to compare how traditional written refuges translate to the modern online forums of catharsis to examine how exile remains a creative force that transcends time and form. Recent exilic works such as those of Isabel Allende or Rosa Yassin Hassan may offer stimulating material for the examination of this continued lingual asylum. Matters of exile create a thread of communal understanding that unites peoples of infinitely distinct backgrounds in a literary sanctuary.

Works Cited

"Augustus." Historic World Leaders, edited by Anne Commire, Gale, 1994. Gale In Context: World History.

Fredericks, B. R. “Tristia 4.10: Poet’s Autobiography and Poetic Autobiography.” Transactions of the American Philological Association (1974-), vol. 106, 1976, pp. 139–54.

GREBE, SABINE. “Why Did Ovid Associate His Exile with a Living Death?” The Classical World, vol. 103, no. 4, 2010, pp. 491–509.

Levin, Donald N. “Phèdre and Oenone.” Donald N. Levin Classics Papers, 1947–1991, MS 46, Woodson Research Center, Fondren Library, Rice University, pp. 51–68. Rice University Repository.

Pavel, Thomas. “Exile as Romance and as Tragedy.” Poetics Today, vol. 17, no. 3, 1996, pp. 305–15.

Olsson, Anders. "Exile and Literary Modernism". Modernism, edited by Astradur Eysteinsson, Vivian Liska, Anke Brouwers, Vanessa Joosen, Nathan Van Camp, Dirk Van Hulle, Katrien Vloeberghs and Björn Thor Vilhjálmsson, John Benjamins Publishing Company, 2007, pp. 735-
754.

Ovid, and Peter Green. The Poems of Exile: Tristia and the Black Sea Letters. U of California P, 2005.

Racine, Jean, and Margaret Rawlings. Phedre. Bordas, 1965.

Rissanen, Matti. “THE THEME OF ‘EXILE’ IN ‘THE WIFE’S LAMENT.’” Neuphilologische Mitteilungen, vol. 70, No. 1, 1969, pp. 90–104.

Tobin, Ronald W. "Jean Racine". Encyclopedia Britannica, 17 Apr. 2025.

The Effects of Sleep, Stress, and Screen Time on Visual Episodic Memory in Adolescents
By Eileen Chen

Abstract

Visual episodic memory refers to one's ability to remember the images and visuals of past experiences. Protecting one's visual episodic memory is important to support learning, problem-solving, and overall cognitive health in adolescence. Today's teens are living in a "Digital Age" marked by increasing screen usage and stress levels, and decreasing sleep, experiences which negatively affect brain structures and may possibly influence visual episodic memory. To further understand these connections, we designed and conducted a survey-based study, which explored the effects of sleep, stress, screen time, and social media use on visual episodic memory in 16 adolescents. This study was designed in such a way that most of the participants who enrolled completed the study, indicating that this study's approach is feasible and could be applied to future studies. However, in this study, no significant relationships were found between sleep, stress, screen time, and social media use on the percent of correct answers or reaction time on a visual episodic memory test (developed by CogniFit). This study has potential limitations, such as a small sample size and the influence of confounding variables, highlighting the complexity of isolating variables that affect cognition. More studies are needed to confirm if there is a lack of correlation.

Introduction

Every day, humans process a stream of images, some of which are quickly forgotten, while others are stored and later recalled. Visual episodic memory, a type of long-term memory, refers to the ability to recall images and visuals that were seen in the past, which can persist for years to even lifetimes (Schurgin and Flombaum, 2018). This differs from working memory, a type of short-term memory, which refers to remembering information and manipulating it to perform complex tasks like problem-solving, and is limited to remembering about 3-5 "chunks" of information for seconds to tens of seconds (Cowan, 2015; Cowan, 2010; Leroy et al., 2024).

Episodic memory is primarily controlled by the medial temporal lobe in the brain, including the hippocampus, which plays a crucial role in encoding and retrieving information from past events (Dickerson and Eichenbaum, 2009). Previous research suggests that numerous factors influence episodic memory, including the time between encounters, age, and diet, due to their effects on brain regions, such as the hippocampus (Schurgin and Flombaum, 2018; Kinugawa et al., 2013; Cansino et al., 2021).

Although the effects of sleep and stress on memory have been extensively studied, their effects on visual episodic memory remain areas of ongoing investigation. In addition, screen time is a relatively new variable of interest. In 2023, American adolescents spent an average of 6-9 hours on media-related activities each day, considerably higher than the recommended 2 hours (Santos et al., 2023). Moreover, in 2018, 97 percent of Americans aged 13 through 17 used at least one of the following social networking sites: YouTube, Instagram, Snapchat, Facebook,

Twitter, Tumblr, and Reddit (Vidal et al., 2020). Excessive screen time is linked to increased risk of mental health disorders (e.g., depression and anxiety) and sleep issues, both of which may affect visual episodic memory performance (Nakshine et al., 2022). Understanding how these factors may impact visual episodic memory is crucial in order to protect one's ability to recall previous experiences, which can aid in learning, problem-solving, and overall cognitive health, improving youth learning and mental health.

The purpose of this experiment is to examine how sleep duration, perceived stress, and daily screen time relate to visual episodic memory performance. Given evidence that lifestyle and environmental factors can influence cognitive functioning, this study seeks to better understand how these commonly experienced variables may affect both memory accuracy and processing speed. To address this question, participants completed self-report measures assessing sleep, stress, and screen use, as well as a standardized visual episodic memory task. Performance was evaluated using accuracy and reaction time as primary outcome measures.

Methods and Materials

A 9-question Qualtrics survey was created to assess participants' daily sleep, stress, and screen time use. The survey takes about 5 minutes to complete and asks participants to report the number of hours of sleep from the previous night, total screen time for the day, and amount of time on social media using short-answer text boxes. Participants were also asked to rate their current stress level from 1 to 10 (1 being the lowest, and 10 the highest). In addition, participants were asked to provide basic demographic information, including name, age, and gender. Participants were invited to take part in the study via direct outreach and instructed to complete the survey 2-3 times per week over a two-week period. They received reminder emails prompting completion. Informed consent was obtained, and approval was granted by an Institutional Review Board (IRB) consisting of a pediatrician, a high school biology educator, and a high school vice principal. Because this is an observational study, the review was expedited. Responses were linked to participants' legal names to allow for data tracking longitudinally (across the 2 weeks the study took place). To protect participants' data, all data was stored in a password-protected Qualtrics form on one password-protected computer.

After completing the questionnaire each time, participants were directed to CogniFit's free online visual episodic memory test, in which images appear sequentially, and respondents press the spacebar on their keyboard whenever an image repeats (Cognifit; Figure 1). Participants then recorded their percent correct and reaction time in the survey. A total of 17 participants completed the survey 66 times, and the Cognifit online visual episodic memory test was taken 66 times. Those who completed fewer than three survey entries were excluded from analyses, resulting in the removal of one participant (n=1 number of participants excluded), resulting in a final sample of 16 participants, whose 64 survey/test entries were analyzed.

Figure 1: CogniFit's Visual Episodic Memory Test

Image 1 Image 2 Image 3

In CogniFit's Visual Episodic Memory Test, images are shown one at a time to participants. Participants are prompted to press the spacebar on their keyboard when seeing an image that has been shown before in the test. It would be correct to not press the spacebar on Image 2 but press the spacebar on Image 3.

Results

Table 1: Demographic Data of Final Sample of Participants

Number of Participants	16
Number of Responses	64
Gender (m:f)	1:15
Age (mean, median, SD, range)	17.1, 17, 1.0, 4

Table 2: Summary of Results

	Mean	Median	Standard Deviation	Range
Sleep in Hours	7.7	8	2	12
Stress (1-10)	6.9	7	2.1	8
Screen Time in Hours	3	3	1.9	9
Time on Social Media in Hours	1.8	1.45	1.6	8
Percent Correct	97.8	98	2.6	9
Reaction Time in Seconds	0.7	0.68	0.1	0.5

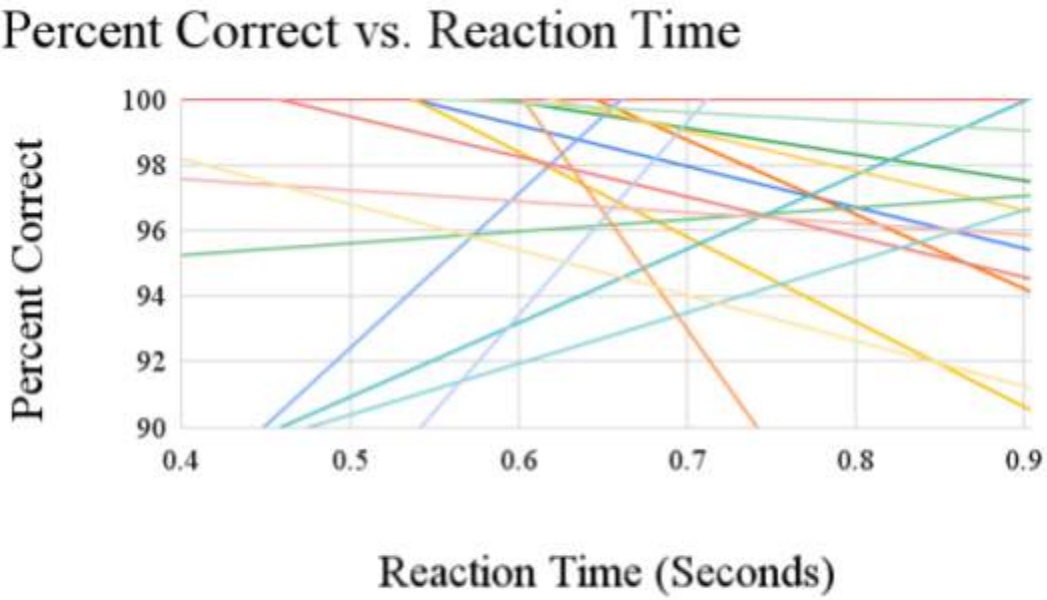


Figure 2. Percent Correct vs. Reaction Time. Each participant's percent correct on CogniFit's Visual Episodic Memory Test was plotted against their reaction time in seconds, and their correlation line was graphed. Each line represents one individual.

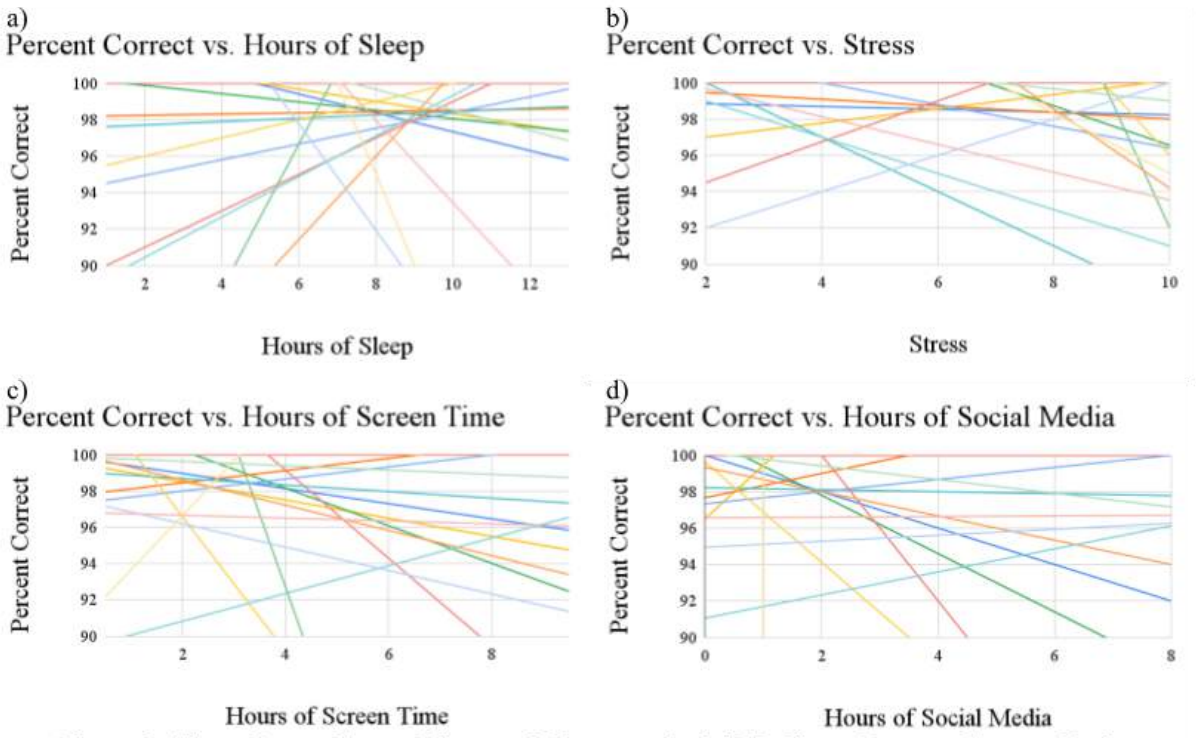


Figure 3: Sleep, Stress, Screen Time, and Hours on Social Media vs. Percent Correct. Each participant's percent correct on CogniFit's Visual Episodic Memory Test was plotted against their (a) number of hours of sleep, (b) amount of stress, (c) hours of screen time, and (d) hours of time spent on social media. Then their correlation line was graphed. Each line represents one individual, illustrating within-individual relationships between the four factors and reaction time across repeated measurements. (a) The relationship between sleep and percent correct was not statistically significant (mean correlation coefficient=-0.002, p=0.990). Similarly, (b) the relationship between percent correct and stress (mean correlation coefficient=-0.118, p=0.408), (c) hours of screen time (mean correlation coefficient=-0.253, p=0.121), and (d) hours of time spent on social media (mean correlation coefficient=-0.119, p=0.401) was not statistically significant. P-values were generated from a one sample *t*-test of the correlation coefficients against a null hypothesis of 0. A p-value of greater than p=0.08 indicates the results were not statistically significant. Overall, across all four variables, it was found that their relationship with percent correct is not statistically significant.

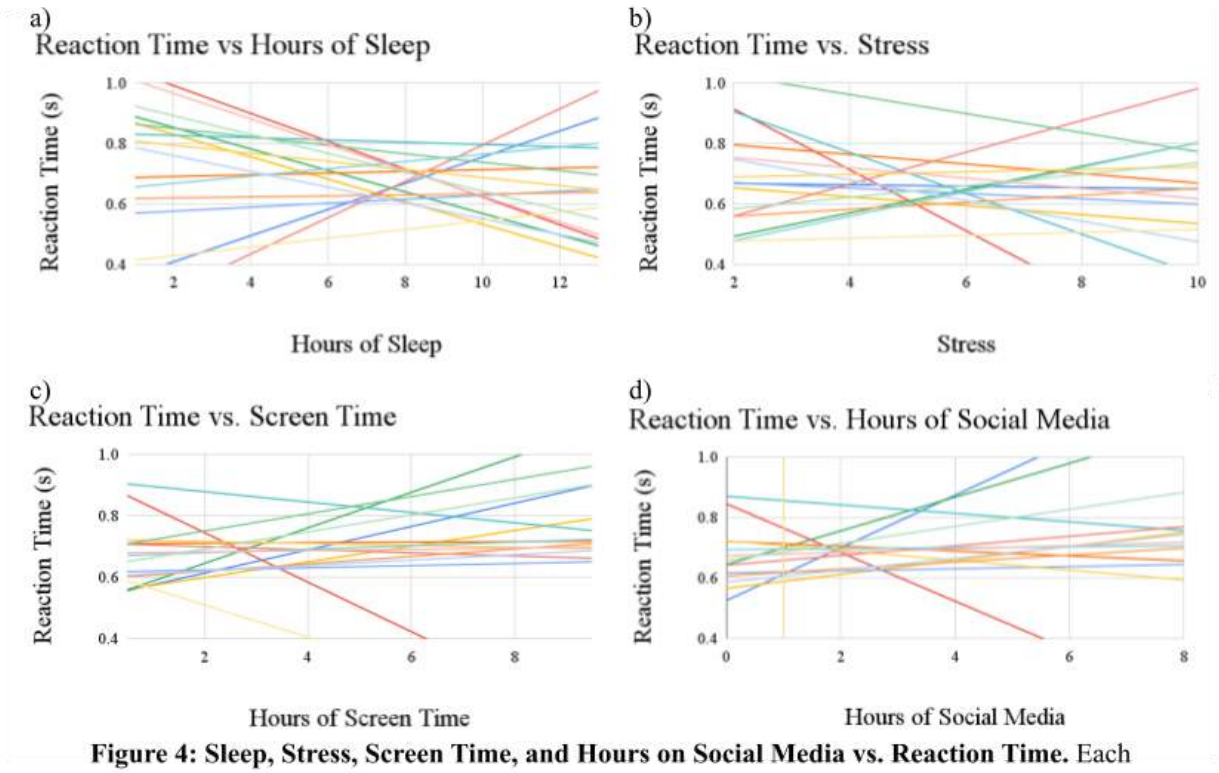


Figure 4: Sleep, Stress, Screen Time, and Hours on Social Media vs. Reaction Time. Each participant's reaction time on CogniFit's Visual Episodic Memory Test was plotted against their (a) number of hours of sleep, (b) amount of stress, (c) hours of screen time, and (d) hours of time spent on social media. Then their correlation line was graphed. Each line represents one individual, illustrating within-individual relationships between the four factors and percent correct across repeated measurements. (a) The correlation between sleep (mean correlation coefficient=-0.090, p=0.540), (b) stress (mean correlation coefficient=-0.131, p=0.371), (c) screen time (mean correlation coefficient=0.192, p=0.200), and (d) hours on social media (mean correlation coefficient=0.168, p=0.176) showed no significant relationship with reaction time. All average correlation coefficients were small, indicating weak linear relationships. P-values were generated from a one sample *t*-test of the correlation coefficients against a null hypothesis of 0. A p-value of greater than p=0.08 indicates the results were not statistically significant. Overall, across all four variables, it was found that their relationship with reaction time is not statistically significant.

A total of 64 responses from 16 participants were analyzed. The sample was heavily skewed towards high school females and had a narrow age range (ages 16-20). The mean percent correct was 97.8% with a standard deviation of 2.6%, and the mean reaction time was 0.7 seconds with a standard deviation of 0.1 seconds (Table 2). The mean amount of sleep was 7.7 hours with a standard deviation of 2.1 hours, the mean amount of screen time was 3.0 hours with a standard deviation of 1.9 hours, and the mean time spent on social media was 1.8 hours with a standard deviation of 1.6 hours (Table 2).

For sleep vs. percent correct, the participants' correlations largely varied in direction, suggesting high variability across participants rather than a consistent relationship (mean correlation coefficient=-0.002, p=0.990; Figure 3a). Similarly, correlations for stress and percent

correct showed mixed directions with no consistent pattern (mean correlation coefficient=-0.118, p=0.408; Figure 3b). For screen time vs percent correct, the correlations also differed in direction between individuals (mean correlation coefficient=-0.253, p=0.121; Figure 3c). Correlations between hours on social media and percent correct showed similar patterns of mixed correlation directions (mean correlation coefficient=-0.119, p=0.401; Figure 3d). Sleep, stress, screen time, and hours on social media all showed mixed correlation directions with percent correct, and none reached significance.

For sleep vs. reaction time, the participants' correlations largely varied in direction, suggesting high variability among individuals and no consistent relationship (mean correlation coefficient=-0.090, p=0.540; Figure 4a). Similarly, correlations between stress and reaction time showed mixed directions across participants (mean correlation coefficient=-0.131, p=0.371; Figure 4b). For screen time vs. reaction time, the correlations showed some variability in direction between individuals (mean correlation coefficient=0.192, p=0.200; Figure 4c). Correlations between hours on social media and reaction time showed similar patterns of mixed correlation directions (mean correlation coefficient=0.168, p=0.176; Figure 4d). Sleep, stress, screen time, and hours on social media all showed mixed correlation directions with reaction time, and none of the relationships were statistically significant.

Discussion

The study demonstrated that all average correlation coefficients were small, indicating weak linear relationships. Overall, none of the factors studied (sleep, stress, screen time, and time on social media) showed statistically significant relationships with percent correct or reaction time.

However, it must be noted that this study has potential limitations. There was a small sample size of 16 individuals with a similar age range (16-20 years) and heavily skewed towards females, making the sample not representative of all youth. Furthermore, there was a lack of variation in some individuals, such as getting 100 percent correct on all visual episodic memory tests taken or getting the same amount of sleep across all days, reducing the reliability of the calculated correlation coefficients. Repeatedly taking the visual episodic memory test may have also caused participants to improve their accuracy with more practice. Furthermore, participants had an average accuracy of 97.8% correct. This ceiling effect likely reduced variability in scores, limiting the ability to detect within-person associations. Future studies could address this by conducting more challenging memory tasks, such as increasing the number of items or increasing retention time. Various confounding variables may have also influenced the data. For instance, the time of day each memory test was taken differed across individuals, possibly influencing the data based on how well people focus during a certain time of day. Diet or learning disabilities, due to their effect on regions in the brain, may have also played a role.

Most of the participants who enrolled completed the study, indicating that this study's approach is feasible and could be used for further in-depth studies. Future research could improve this study by using a larger sample that is more representative of the population of

youth. The survey could also run for more time and have each individual take more tests to allow for more accurate comparisons between factors and their effects on visual episodic memory.

Conclusion

The purpose of this experiment is to study the effects of sleep, stress, and screen time on visual episodic memory. Most of the participants who enrolled completed the study, indicating that further in-depth studies could be conducted using similar methods. No significant relationships were found between sleep, stress, screen time, and hours of social media on percent correct or reaction time in CogniFit's online visual episodic memory test, possibly due to limitations in the study like small sample size and the presence of confounding variables.

In today's Digital Era, where long screen times, sleep deprivation, and high stress are prevalent among teens, understanding how factors such as sleep, stress, screen time, and time on social media may impact visual episodic memory is crucial to protecting one's ability to recall past experiences. Although this study found no significant relationships, future research should be conducted to better understand how these factors influence visual episodic memory, which can promote learning and help an individual support their overall brain health.

Acknowledgements

I express my deepest gratitude to those whose support and advice have been invaluable for the completion of this research endeavor: Dr. Joshua Lange, for teaching me how to conduct my own research, and Dr. Poonam Agarwal, for guiding me through the publication process.

Conflict of Interests

The author declares that there are no conflicts of interest related to this work.

Works Cited

Cansino, Selene et al. “Influence of Dietary Nutrient Intake on Episodic Memory Across the Adult Life Span.” Frontiers in aging neuroscience vol. 13 724595. 30 Aug. 2021, doi:10.3389/fnagi.2021.724595. Accessed 2 Nov. 2025.

“CogniFit.” CogniFit Brain Fitness - Try It Now!, www.cognifit.com/aplicaciones/html5/public/assessment/ASSESSMENT~@~VISUAL_EPISODIC?testButtonUrl=www.cognifit.com/memory-test%3Freg=true%26regut=normal%26irclickid=VzqWpBX0XxyNR4K0-Lw1QVHwUkAQs7V9Xx8OzQ0%26irgwc=1. Accessed 2 Nov. 2025.

Cowan, Nelson. “Working Memory Underpins Cognitive Development, Learning, and Education.” Educational psychology review vol. 26,2 (2014): 197-223. doi:10.1007/s10648-013-9246-y. Accessed 5 Jan. 2026.

Cowan, Nelson. “The Magical Mystery Four: How is Working Memory Capacity Limited, and Why?.” Current directions in psychological science vol. 19,1 (2010): 51-57. doi:10.1177/0963721409359277. Accessed 5 Jan. 2026.

Dickerson, Bradford C, and Howard Eichenbaum. “The episodic memory system: neurocircuitry and disorders.” Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology vol. 35,1 (2010): 86-104. doi:10.1038/npp.2009.126. Accessed 12 Dec. 2025.

Kinugawa, Kiyoka et al. “Aging-related episodic memory decline: are emotions the key?.” Frontiers in behavioral neuroscience vol. 7 2. 1 Feb. 2013, doi:10.3389/fnbeh.2013.00002. Accessed 12 Dec. 2025.

Leroy, Nathan et al. “Working memory capacity for continuous events: The root of temporal compression in episodic memory?.” Cognition vol. 247 (2024): 105789. doi:10.1016/j.cognition.2024.105789. Accessed 21 Dec. 2025.

Nakshine, Vaishnavi S et al. “Increased Screen Time as a Cause of Declining Physical, Psychological Health, and Sleep Patterns: A Literary Review.” Cureus vol. 14,10 e30051. 8 Oct. 2022, doi:10.7759/cureus.30051. Accessed 7 Nov. 2025.

Santos, Renata Maria Silva et al. “The associations between screen time and mental health in adolescents: a systematic review.” BMC psychology vol. 11,1 127. 20 Apr. 2023, doi:10.1186/s40359-023-01166-7. Accessed 2 Nov. 2025.

Schurgin, Mark W, and Jonathan I Flombaum. “Properties of visual episodic memory following repeated encounters with objects.” Learning & memory (Cold Spring Harbor, N.Y.) vol. 25,7 309-316. 15 Jun. 2018, doi:10.1101/lm.047167.117. Accessed 7 Nov. 2025.

Vidal, Carol et al. “Social media use and depression in adolescents: a scoping review.” International review of psychiatry (Abingdon, England) vol. 32,3 (2020): 235-253. doi:10.1080/09540261.2020.1720623. Accessed 7 Nov. 2025.

Early-Season Crop Yield Prediction Using Satellite-Derived Vegetation Indices and Attention-Based LSTM Models By Kritin Nori

Abstract

Accurate end-of-season crop yield prediction is crucial for global food security and agricultural planning. Without it, farmers risk over- or under-investing in resources which risks profit loss. Previous research has been limited to the use of specific crops and regions, or the use of traditional machine learning models that do not capture complex patterns, affecting scalability, generalizability, and real-world applicability. This study proposes the use of a time-series deep learning model, Long Short-Term Memory(LSTM), to predict end-of-season crop yields utilizing free remote sensing data(MODIS) along with satellite based vegetation indices such as Normalized Difference Vegetation Index(NDVI) and Enhanced Vegetation Index(EVI) both of which measure the greenness and health of the crops in a given plot of land. We used a total of four crops(Yams, Sorghum, Millets, and Maize) ranging from countries such as Ghana to Nigeria and Cameroon to Ethiopia. In doing this, we not only have a variety of datasets to extract data from, but we also use modern machine learning models which helps to analyze complex data patterns. The LSTM model was compared to a baseline model to prove that deep learning models are more accurate than traditional models and the results back that up: the LSTM model had an R^2 score of 0.989 meaning that the predictions nearly mirrored real-life outcomes, while the baseline model had an R^2 score of 0.838 relevant enough, but underperforming the LSTM model by a wide margin. From these results, it's proven that deep learning models like LSTM perform well in accurately predicting end-of-season crop yields, helping farmers prepare to invest in resources while maintaining accuracy.

Keywords LSTM, Deep Learning, Satellite-Based, NDVI, EVI

Introduction

With modern technologies ever-evolving and taking over the world, agriculture still stands as one of the largest sectors for mankind. It is because of the many farmers that the world is able to sustain, but these farmers and the agricultural sector as a whole are facing a multitude of problems such as climate variability, various environmental stressors such as droughts, floods, pollution, and soil degradation. With these issues increasing at an alarming rate, global food security and agricultural planning of resources are at risk.

Wang et al. (2018) have focused their research primarily on a singular crop, soybeans, and implemented transfer learning with deep learning models such as convolutional neural networks(CNNs) and Transformers. The study also paired the models with satellite-based images as inputs to measure how well the crops are expected to yield. Even as the idea enhances knowledge on what is known in the field of agricultural sciences, its scope is limited to a singular region, Argentina, and its implementation of just soybeans makes generalizing the data a bit difficult to more larger regions. Judith et al.(2025) used NDVI features paired with satellite

images to forecast yields and achieved an accuracy and recall over 90% for most classes, with classes meaning different classifications of crop health(healthy or unhealthy). That said, this study uses a fully connected neural network or FCNN, which is a classical machine learning model, and the study itself mentions not using a deep learning model like CNNs or Transformers, as a limitation for accurate yield prediction. Several other studies such as Joshi et al. (2024) and Mahalakshmi et al. (2025) implement deep learning models such as CNNs and other deep learning models like AdaBoost. That said, both studies performed their testing on limited data, meaning that the models may be biased and as a result not scalable to other crops or regions. The typical base assumption in this field is that deep learning models generally perform well in providing accurate forecasts of crop yields, outperforming traditional methods of gathering data like field observation or interviews with farmers Mahalakshmi et al. (2025).

This study builds on the foundation set by previous researchers. In order to combat a lack of generalizability, scalability and accuracy, this study proposes the use of a deep-learning time series model called Long Short-Term memory or LSTM for accuracy. In order to be able to scale the data to larger regions, this study proposes the use of a multitude of crops across four different countries for more precise results. By doing this, we not only have enough data, a limitation of past studies, but we also have a deep learning model capable of remembering inputs from the past and use that information to predict future trends.

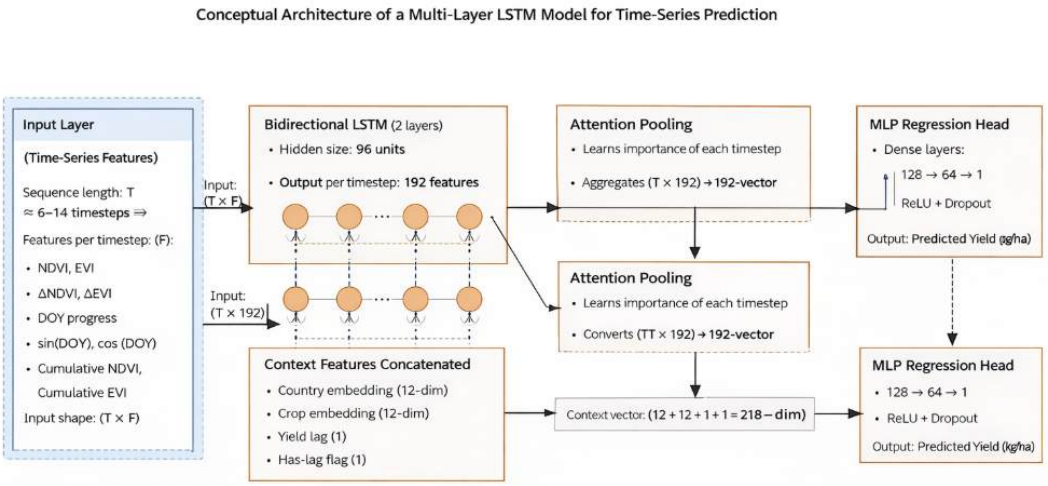


Figure 1: Architecture of the attention-based LSTM model used in this study

Time-series inputs consisting of NDVI, EVI, and derived seasonal features (T * F) are passed through a two-layer bidirectional LSTM (hidden size = 96 per direction). The resulting sequence representations are aggregated using an attention pooling mechanism that learns the relative importance of different timesteps. The pooled representation is concatenated with country embeddings, crop embeddings, and a normalized yield-lag feature before being passed through a multi-layer perceptron (128 to 64 to 1) to produce the final yield prediction.

1. Methodology

1.1 Study Area and Crops

The primary focus of this experiment is yield prediction throughout four countries: Kenya, Ethiopia, Ghana, and Nigeria. We focused on these countries specifically because they rely heavily on agriculture for their economy and food security, which increases the real-world applicability of this study. We analyzed four staple crops from these countries: cassava, millets, sorghum, and yams. These crops were chosen because they are the key crops that support food security in these sub-Saharan countries and all of Africa.All the analysis was conducted at the first administrative level (ADM1), relating to states or provinces. ADM1 boundaries were obtained from the geoBoundaries database. This ADM1 level was chosen because it provides sufficient information by region while still having reliable data available.

1.2 Satellite Data and Vegetation Indices

Vegetation data from satellites was derived from the MODIS MOD13Q1 dataset, which offers 16-day composite values for both the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) on a 250-m spatial resolution (each pixel represents 250 by 250 dimensions of land). The dataset was extracted via the Google Earth engine platform. This dataset is used for agricultural mapping as it is capable of detecting the seasonal growth cycle of agricultural crops in each plot of land.

1.3 Vegetation Data

Satellite-based data is broad, meaning that it will provide data throughout vast regions rather than smaller plots of land. Smaller plots of land helped us focus exclusively on crops rather than other vegetation. Spatial Productive Allocation Model (SPAM) 2017 dataset, which is provided in raster(GEOTIFF) format, was used to identify specific places where crops were grown in larger regions. Raster format shows us the plots of land with small pixels.

For each crop, the corresponding SPAM raster layer was combined with ADM1 boundary polygons(outlines of plots of land) using geospatial operations. For each ADM1 region and crop, zonal statistics were computed to estimate total harvested area. These harvested areas were then used as spatial weights when aggregating NDVI and EVI values across all ADM1 units. The final crop-weighted vegetation index for each country-crop-year- Day of Year (DOY) combination was calculated as a weighted average, where regions with larger harvested areas contributed more strongly to the final signal. This ensured the data reflected how the crops actually grew, not just general vegetation.

1.4 Yield Data

Crop yield data was used from a publicly available dataset from the Food and Agricultural Organization website. The dataset consisted of crop data by year, country, and the yield value of the crop in the corresponding year. Yield represents agricultural productivity per unit area, typically expressed in kilograms per hectare (kg/ha). Yield data was cleaned and preprocessed to remove any unnecessary or missing information.

1.5 Time-Series Construction and Feature Engineering

The data was split into a combination of country, crop, and year. Meaning that each crop was paired with a year, and a country. For each country-crop-year combination, a time-series was created until the 209 early-season cutoff period. This emulates real world scenarios, as we are using early season data to forecast end-of-season data.
Several additional features were derived from the raw NDVI and EVI signals to enhance model learning:

- First-order differences of NDVI and EVI were computed to capture short-term vegetation change.
- Cumulative mean values were computed to represent progressive biomass accumulation throughout the season.
- Sinusoidal transformations of DOY (sine and cosine) were used to encode seasonal timing.
- A normalized seasonal progress feature (DOY divided by 365) was included.

Missing satellite values were estimated using nearby points, and all features were scaled using only the training data to prevent the model from using future information(data leakage).

1.6 Model Architecture

The model used in the experiment is Long Short-Term Memory or LSTM. LSTM is a deep learning time-series model that is capable of capturing inputs and remembering those inputs to make accurate predictions on unseen data, which is why it was chosen. The model will take time-series data as input and produce one predicted value(yield) per country, crop, and year.

Each sequence included NDVI/EVI numbers adjusted using crop-specific weights pulled from MODIS information. Alongside that, added traits helped reflect how crops evolved across time. At each moment, nine elements made up the set: raw NDVI/EVI levels, shifts between those readings, markers tied to when seasons unfolded - shown via day-of-year counts - and paired sine-cosine patterns. How long each string lasted depended on where the season ended. Since they weren't uniform, these shifting spans were managed using packed training methods.

A setup using two layers of bidirectional LSTM formed the core, each direction holding 96 hidden units, along with an attention-based pooling step. This attention mechanism highlights key moments across a growing season that influence predictions, collapsing the full timeline into one compact summary vector. Specifically, the model reads the NDVI and EVI inputs and learns the crop pattern in a time-series manner. Time-series data is data stored chronologically by date, starting from day 1 till day 209 in 16-day intervals. Once the model understands the crop growth patterns, it distinguishes the differences in crop growth. Instead of treating Ghana cassava and Ethiopia sorghum as identical, the model will treat them as different conditions of crop growth as the growth of different crops varies, especially in different countries. The model also uses the lagging technique, which is when it factors in data from previous years. Finally, all this

processing is merged and passed through multiple neural layers that gradually decrease from 218 to 128 to 64 until one final number remains which is the predicted yield.

1.7 Model Training, Evaluation, and Preprocessing

The dataset was split into training, validating, and test dataset. All but the last four years were used for training the model. The last three years were used for testing the model, and the lone fourth last year was used for validating the testing results. This split based on time intervals ensured no data leakage, meaning the model had no access to future data while training which would have skewed the results.

The model was trained using the Adam optimizer and Smooth L1 loss function. Both of these tools helped the model understand how wrong its decisions were during training and helped the model adjust its predictions. We applied early stopping, which was when the models stopped making better decisions on data it had not seen before. Model performance was evaluated using three metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R^2 score, each with its own purpose.

To better understand model performance, a baseline model was also implemented. This baseline model received the same treatment as the LSTM, with the exception that it is not as advanced as the LSTM. The main purpose of this baseline was to compare it against the LSTM model and show that the LSTM model performs better in yield predictions. The modeling pipeline used in study is represented by Figure 2.

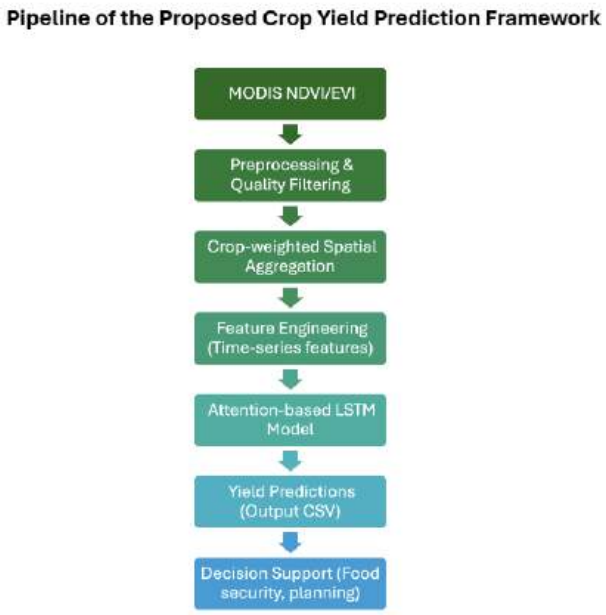


Figure 2: Pipeline of the proposed crop yield prediction framework. Satellite-derived NDVI/EVI time-series data are processed through preprocessing, crop-weighted spatial aggregation, and feature engineering before being modeled using an attention-based LSTM to generate early-season yield predictions for decision support.

2. Experiments and Results

2.1 Experimental Design and Evaluation Strategy

The main objective of this study was to determine whether early-season satellite-derived NDVI and EVI values are sufficient enough to forecast end-of-season crop yields by the LSTM model at the regional level in multiple countries. The main idea was to use the earlier years(all but the last four) in the dataset for training. We used the fourth last year for validation, and the last three years for testing. It is also worth mentioning that for the sake of this experiment, we deliberately took the first 209 days, in groups of 16 provided by the dataset, and removed all days after that and classified it as early-season, since anything beyond that would skew accuracy.

The model performance was evaluated using multiple evaluation metrics to provide a summary of how accurate the LSTM model was in forecasting yield. Root Mean Squared Error (RMSE) was used to calculate how far on average the predicted yield value was from the actual yield value, penalizing larger differences more. Mean Absolute Error(MAE) also measures on average how far the predicted was from the actual, but treats all values equally as a difference. For RMSE and MAE, the lower the value, the better the model performed. R^2 score measures how accurate the predictions were. If a R^2 score of a model is closer to one, then it means that the model made near perfect decisions. A score closer to zero would indicate that the prediction of the model was equivalent to making an average guess. The LSTM model used in this study achieved an R^2 score 0.989.

2.2 Dataset Coverage and Preprocessing Outcomes

Table 1 summarizes the datasets used in this study to train and test the model. The MODIS MOD13Q1 dataset provided satellite-based vegetation information, which is the main input for this research. It included the country, ADM1 region name, year, day of year, and NDVI and EVI values recorded throughout each growing season. The FAOSTAT dataset provided the crop types (yams, cassava, sorghum, and millets) and the actual yield values for each crop. The SPAM dataset provided estimates of where each crop is grown within each region, which allowed us to focus the satellite signals on cropland areas. These crops were selected because they are important staple crops in the sub-Saharan African countries included in this study.

Table 1: Summary of datasets used in this study

Country	Crop	Years Covered	ADM1 Units	Temporal Resolution	Data Source
Cameroon	Cassava	2000-2023	~10	16-day	MODIS MOD13Q1, FAOSTAT
Cameroon	Sorghum	2000-2023	~10	16-day	MODIS MOD13Q1, FAOSTAT
Cameroon	Yams	2000-2023	~10	16-day	MODIS MOD13Q1, FAOSTAT
Cameroon	Millets	2000-2023	~10	16-day	MODIS MOD13Q1, FAOSTAT

Ethiopia	All crops	2000-2023	~11	16-day	MODIS MOD13Q1, FAOSTAT
Ghana	All crops	2000-2023	~16	16-day	MODIS MOD13Q1, FAOSTAT
Nigeria	All crops	2000-2023	~37	16-day	MODIS MOD13Q1, FAOSTAT

2.3 Results & Discussion

Table 2 shows the performance of the LSTM model compared to the baseline model. The LSTM model achieved an RMSE value of 713.10, MAE value of 432.81, and an R^2 score of 0.989, significantly outperforming the baseline model, which achieved an R^2 of 0.838.

Table 2: Test-set performance of the crop-weighted within-season LSTM compared to a mean-yield baseline (DOY ≤ 209).

Model	RMSE	MAE	R^2
Mean yield baseline	2765.41	1634.79	0.838
Crop-weighted LSTM (attention)	713.10	432.81	0.989

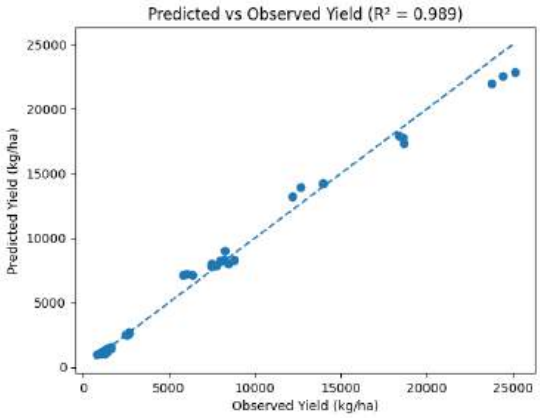


Figure 3: Predicted vs. Observed yield for LSTM Model

The tight clustering shown in figure 5, shows the predicted values are very close to the actual values, confirming the results displayed in Table 2. This shows the LSTM model was able to capture the complex patterns and accurately forecast yields.

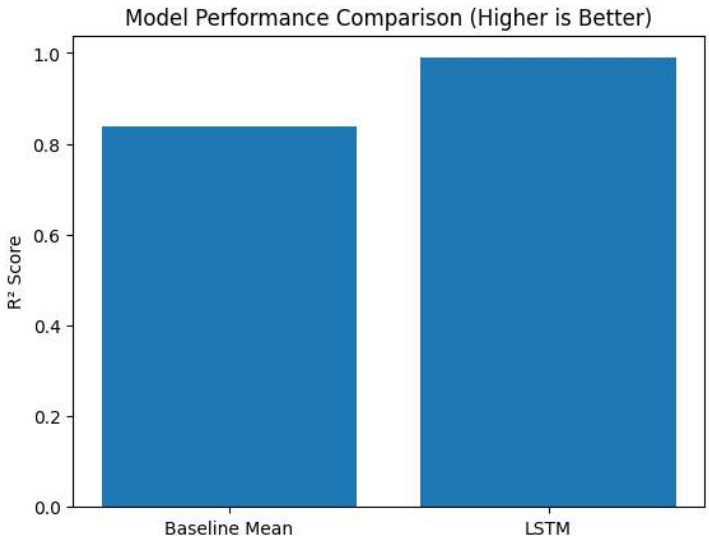


Figure 4: Model Accuracies

This demonstrates that the proposed framework provides meaningful improvement beyond historical averages.

2.4 Error Analysis and Crop-Level Behavior

Table 3 shows information on each configuration that helped achieve model accuracy and why it was so important. The model performed worse when we gave importance to all vegetation equally because the crop signal was mixed with non-cropland, meaning the model factored in trees and plants rather than solely focusing on the desired crops. With crop weighting, we saw a massive difference, as now the model gave importance to primarily crops leading to higher accuracy. When we used NDVI only, the model performance was moderate, since the model was missing inputs that EVI would normally provide. Of course, model accuracy improved after we added both input features, as now the model is provided with full data. With no attention mechanism, the model performed poorly as it weighted each time period equally. Mid-season crop yields are more valuable than early-season crop yields, yet the model treats both values equal, hence the performance drop. Without using the previous year data, the model couldn't train very well as opposed to when it did have all years available, another factor for performance dropoff. It was a combination of all the configurations that helped the model reach optimal accuracy.

Table 3: Effect of model components on prediction performance

Configuration	Description	Performance Impact
No crop weighting	NDVI/EVI averaged uniformly across ADM1 regions	Worse performance due to dilution of crop-specific signal
Crop-weighted NDVI/EVI	Vegetation indices weighted by harvested crop area	Improved accuracy and stability

NDVI only	Uses NDVI without EVI	Moderate performance
NDVI + EVI	Combines complementary vegetation signals	Improved performance
No attention	Mean pooling across time steps	Reduced ability to focus on critical growth stages
Attention pooling	Learns the importance of different seasonal phases	Best overall performance
No yield lag	Removes previous-year yield information	Lower early-season accuracy

2.5 Summary of Results

Collectively, All three tables show that the modeling pipeline is both accurate and diverse. Table 1 validates the datasets and shows how each dataset was important in the pipeline. Table 2 ensures accurate prediction by the LSTM model, and Table 3 shows us how each configuration played an important role in model accuracy. Together these results support the idea that deep learning time-series models can accurately forecast end-of-season yields given early-season data.

Conclusion

Previous studies in this field have laid the foundation for this one, but typically focused their work on a singular aspect: either using deep learning techniques, focusing on specific crops for yield prediction, or narrowing the scope of their study to a specific region. The gap which we are trying to bridge in this study is to integrate all the three aspects so as to provide accurate prediction to farmers throughout the regions. This was achieved by using free remote sensing data from the MODIS dataset, analyzing several crops across 4 countries (Ghana, Cameroon, Nigeria, Ethiopia) and using input features such as NDVI and EVI. Leveraging all these three we successfully built a deep learning model (LSTM) that achieved an R^2 score of 0.989 out pacing the baseline which has achieved an R^2 score of 0.838, showcasing that deep learning models demonstrate strong performance in yield prediction. This model can be extended into a mobile application that farmers can use efficiently to forecast their season crop yields and effectively use resources. While this study shows accurate model performance, future work can be expanded to have additional crops included in the dataset as well as include a few more countries from other regions, incorporate soil quality and other environmental factors.

Acknowledgement

I would like to thank my mentor Ms. Aashna Saraf for all of her guidance and support throughout this entire process. I would also like to thank my parents and school without whom none of this would have been possible.

Works Cited

Wang, A., Tran, C., Desai, N., Lobell, D., & Stefano Ermon. (2018). Deep Transfer Learning for Crop Yield Prediction with Remote Sensing Data. COMPASS ’18 Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies. https://fse.fsi.stanford.edu/publication/deep-transfer-learning-crop-yield-prediction-remote-sensing-data

Judith, J., Tamilselvi, R., Parisa, B. M., Lakshmi, S. Sathiya Pandiya, Panthakkan, A., Mansoori, S. A., & Ahmad, H. A. (2025). Remote Sensing Based Crop Health Classification Using NDVI and Fully Connected Neural Networks. ArXiv.org. https://arxiv.org/abs/2504.10522

Kogan, F., Salazar, L., & Roytman, L. (2011). Forecasting crop production using satellite-based vegetation health indices in Kansas, USA. International Journal of Remote Sensing, 33(9), 2798–2814. https://doi.org/10.1080/01431161.2011.621464

Mahalakshmi, S., Jose Anand, A., & Partheeban, P. (2025). Soil and crop interaction analysis for yield prediction with satellite imagery and deep learning techniques for the coastal regions. Journal of Environmental Management, 380, 125095. https://doi.org/10.1016/j.jenvman.2025.125095

Runfola, D., Anderson, A., Baier, H., Crittenden, M., Dowker, E., Fuhrig, S., Goodman, S., Grimsley, J., Layko, D., Melville, A., Mulligan, J., Quentin, L., Rogan, J., & Steinberg, D. (2020). geoBoundaries: A global database of political administrative boundaries [Data set]. https://www.geoboundaries.org/

Didan, K. (2021). MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD13Q1.061

International Food Policy Research Institute. (2019). Spatial Production Allocation Model (SPAM) 2017 v2.0 [Data set]. https://doi.org/10.7910/DVN/PRFF8V

Asymmetric Pass-Through in U.S. Credit Card Interest Rates: 2007–2025
By Ian Drummond

Abstract

Although standard monetary transmission theory assumes symmetric pass-through, U.S. credit card APRs remain elevated following Federal Reserve (Fed) rate cuts. This paper applies descriptive time series analysis using Fed and industry data to examine the mechanisms behind this asymmetric pass-through from April 2007–November 2025. The results show that credit card APRs rapidly adjusted upward during the Fed's tightening cycles, but declined only modestly during easing periods, with APRs increasing 4.57% on average following hikes in the effective federal funds rate (EFFR) and only decreasing 0.95% on average following cuts in the EFFR. High concentration in the credit card market, measured by a four-firm concentration ratio (CR4) of 53.03%, suggest that credit card pricing is influenced not only by funding costs but also by market concentration and issuer strategy. These findings highlight the limits of monetary policy transmission in concentrated consumer credit markets.

Introduction

While standard monetary transmission theory assumes symmetric pass-through of policy rates to consumer lending (Mishkin), empirical evidence suggests this theory fails in practice. In particular, Karagiannis et al. documents asymmetric interest rate pass-through in the post-financial crisis period with lending rates exhibiting downward rigidity. Understanding this asymmetry is vital, as it affects how effectively monetary policy translates into changes in borrowing costs for households.

This paper expands upon those insights, examining trends in the U.S. consumer credit market across multiple monetary cycles between 2007 and 2025. I hypothesize that elevated APRs stem from delayed pass-through caused by an oligopolistic market structure that favors non-price competition. By blending time series analysis of average APRs and the EFFR with market concentration and rewards program data, this study examines the structural factors restricting the downward flexibility of consumer lending rates.

Literature Review

Beyond credit risk, industry structure is a major determinant of price rigidity. The U.S. credit card market is characterized by high industry concentration among a few large issuers. Neumark and Sharpe demonstrate that in such concentrated deposit markets, prices become 'sticky' due to oligopolistic coordination and reduced competitive pressure to undercut rivals. Expanding on this insight, Hannan and Berger show that this asymmetry is highlighted in interest rate adjustments, with downward rigidity exceeding upward rigidity. Hannan and Berger attribute this behavior to the oligopolistic pricing behavior exhibited in credit card markets, limiting incentives to reduce APRs following decline in benchmark rates.

Scholnick documents delayed pass-through because of strategic pricing behaviors, which reinforce downward stickiness. Consumer credit markets are characterized by limited borrower attention and high switching costs, which weaken competitive pressure on interest rates, resulting in downward rigidity. Prior research by Agarwal et al. show that differences in borrower type and broader macroeconomic conditions elicit different bank lending rates' responses to monetary policy adjustments.

In oligopolistic markets, competition frequently shifts to non-price based strategies (Siddharthan). Thus, rather than compete by lowering APRs for consumers, major issuers compete via rewards programs, sign-up bonuses, and service features. The strategic focus on non-price competition further reduces downward pressure on APRs, despite cheaper funds following benchmark cuts, due to limited entry and threats from rivals.

Data and Methodology

1. Data Sources

This study employs publicly available data between February 2007 and November 2025, with longer-term market concentration data used for context. The federal funds target rate and the effective federal funds rate (EFFR) are sourced from the Board of Governors of the Federal Reserve System. Average credit card APRs are obtained from the same institution. The credit card market concentration data for 1994, 2016, and 2022 are drawn from
from the Consumer Financial Protection Bureau's (CFPB) 2023 Consumer Credit Card Market Report. The credit card market concentration data, including the four firm concentration ratio (CR4), for 2024 was drawn from the Bank Policy Institute. Finally, data on credit card rewards program expenditures were obtained from IBISWorld's January 2025 report on Card Reward Program Services.

2. Methodology

Given the exploratory nature of this analysis, the methodology applied in this paper is a descriptive time series analysis of patterns over time, designed to document patterns in credit card interest rate behavior across monetary policy regimes, without asserting causal relationships.

The primary instrument for analysis is the interest rate spread between average credit card APRs and the effective federal funds rate, defined as:

$$Spread_t = APR_t - EFFR_t$$

This spread captures the extent to which consumer credit card rates adjust relative to changes in monetary policy rates, avoiding the need for model-based estimates of pass-through.

The sample period is divided into two regimes:

- Tightening phases: December 2015–December 2018; March 2022–August 2023
- Easing phases: August 2007–December 2015; March 2020–March 2022; September 2024–November 2025.

The analysis compares the APR–EFFR spread across these monetary regimes using summary statistics, including mean and median spreads. This approach highlights asymmetries in adjustment without imposing assumptions or estimating coefficients.

Contextual factors are also evaluated. Market structure is assessed using the four-firm concentration ratio (CR4) to contextualize issuers' pricing power. Data on rewards program expenditures were reviewed to assess the prevalence of non-price competition, which may weaken incentives to adjust interest rates downward during easing cycles.

The methodological approach employed in this paper is descriptive and correlational. It does not control for potential confounding variables; therefore, the associations identified should be interpreted as suggestive rather than causal. Furthermore, the APR-EFFR spread analysis focuses on 2007–2025; longer-term CR4 data are used for context. Future research applying econometric analysis to individual issuers could better isolate causal effects and account for issuer-specific factors.

Results and Analysis

1. Timeline of Rates

In response to the Great Recession, the Fed lowered the EFFR from 5.25% in September 2007 to 0% in December 2008. It then remained at the zero lower bound (0–0.25%) until December 2015. Using May 2007 as the baseline, the average interest rate on credit cards at commercial banks fell from 13.46% to 12.03% in November 2008. The average credit card APR increased from 12.03% in November 2008 to a peak of 14.26% in February 2010 before declining to 12.22% by November 2015.

The Fed raised the EFFR from 0-0.25% in December 2015 to 2.25–2.50% in December 2018. The average credit card APR rose from 12.22% in November 2015 to 14.73% in November 2018.

The Fed decreased the EFFR from 2.25-2.50% in July 2019 to 0-0.25% in March 2020, and the EFFR then remained at the zero lower bound (0–0.25%) until March 2022. Using August 2019 as the baseline, credit card APRs decline from 15.10% to 14.52% in May 2020.

From March 2022, to July 2023, the Fed raised the EFFR from 0–0.25% to 5.25–5.50%. Using February 2022 as the baseline, the average interest rate on credit cards at commercial banks increased from 14.56% in February 2022 to 21.19% in August 2023. APRs continued to rise after the tightening cycle ended to a peak of 21.76% in August 2024. The Fed cut rates in September, November, and December of 2024, lowering the target range to 4.25–4.50%, and credit card APRs remain little changed relative to their August 2024 levels.

2. Spread and Pass-Through Analysis

Figure 1 plots the average credit card APR alongside the effective federal funds rate (EFFR) over the period from April 2007–November 2025. During monetary tightening phases, increases in the EFFR are matched by sustained increases in credit card APRs. During the easing phases, declines in the EFFR are not matched by proportional decreases in credit card APRs, which remain elevated.

Figure 2 depicts the average spread between credit card APRs and the EFFR between April 2007–November 2025. During tightening cycles the spread is roughly constant, and during monetary easing cycles, the spread widens.

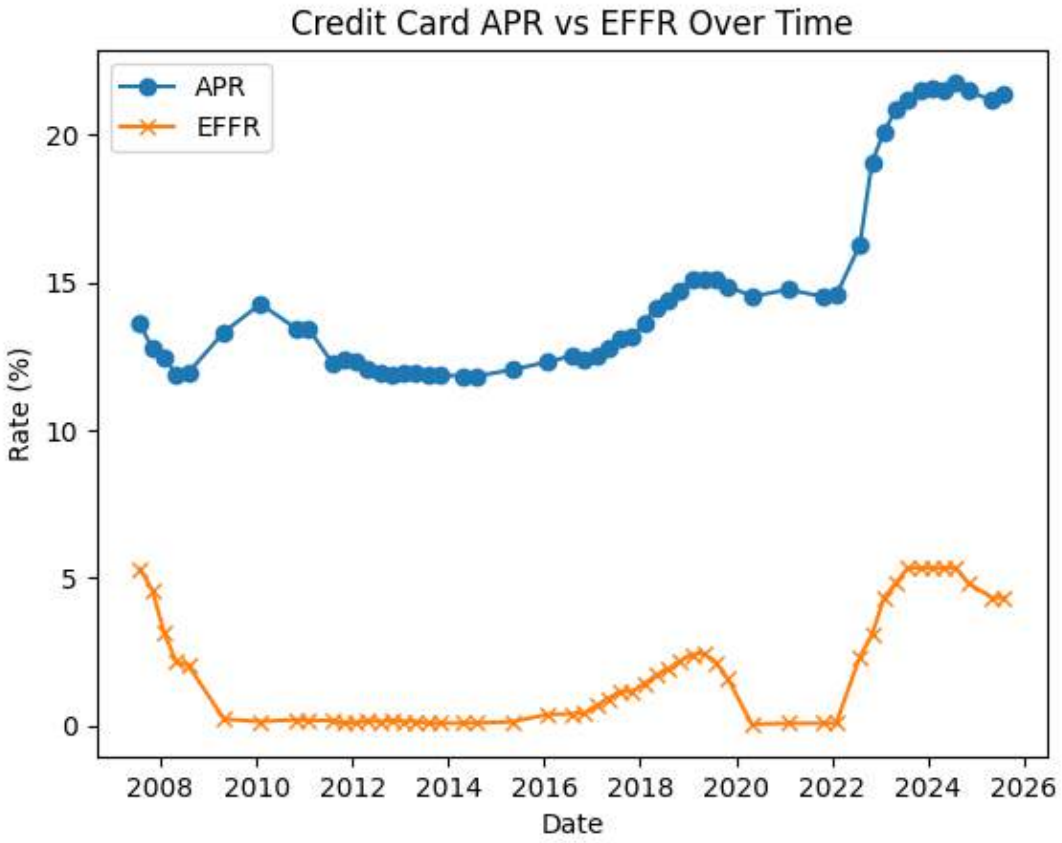


Figure 1: Average Credit Card APR vs. the EFFR between April 2007-November 2025 (Board of Governors of the Federal Reserve System).

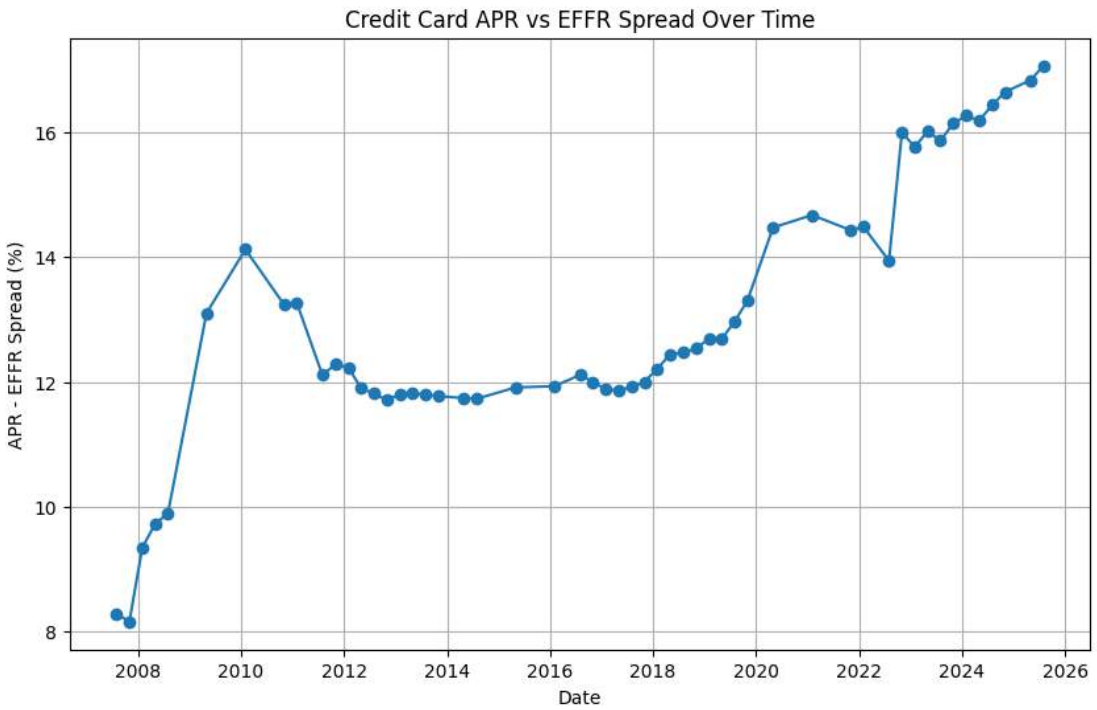


Figure 2: Credit Card APR vs. EFFR spread over time (Federal Reserve Bank of New York).

3. Contextual Factors

To better understand pricing behavior, this paper references industry concentration in the U.S. credit card market. According to Table 1, data aggregated by the Bank Policy Institute indicates that the market is highly concentrated among a small number of large issuers, suggesting limited competitive pressure. The Four-Firm Concentration Ratio (CR4) for this industry is 53.03%.

Company Name	Company Type	Credit Card Loans Oustanding ($ Billions)	Market Share
JPMorgan Chase & Co.	Bank Holding Company	185.79	16.33%
Citigroup Inc.	Bank Holding Company	173.01	15.21%
Capital One Financial Corporation	Bank Holding Company	142.38	12.51%
Bank of America Corporation	Bank Holding Company	102.20	8.98%
Discover Financial Services	Bank Holding Company	102.11	8.97%
American Express Company	Bank Holding Company	98.13	8.62%
Synchrony Financial	Savings & Loan HC	97.04	8.53%
Wells Fargo & Company	Bank Holding Company	53.05	4.66%
Barclays US LLC	Bank Holding Company	31.79	2.79%
U.S. Bancorp	Bank Holding Company	28.71	2.52%
The Goldman Sachs Group, Inc.	Bank Holding Company	19.24	1.69%
Bread Financial	Nonbank	17.90	1.57%
United Services Automobile Association	Savings & Loan HC	16.59	1.46%
TD Group US Holdings LLC	Bank Holding Company	15.31	1.35%
Credit One	Nonbank	12.20	1.07%
First National of Nebraska, Inc.	Bank Holding Company	8.12	0.71%
The PNC Financial Services Group, Inc.	Bank Holding Company	6.23	0.55%
Merrick Bank	Commercial Bank	3.91	0.34%
Truist Financial Corporation	Bank Holding Company	3.49	0.31%
Citizens Financial Group, Inc.	Bank Holding Company	2.09	0.18%
Ally Financial Inc.	Bank Holding Company	1.99	0.17%
Fifth Third Bancorp	Bank Holding Company	1.87	0.16%
Stride Bank, National Association	Commercial Bank	1.59	0.14%
United National Corporation	Bank Holding Company	1.46	0.13%
Regions Financial Corporation	Bank Holding Company	1.34	0.12%
Popular, Inc.	Bank Holding Company	1.14	0.10%
BMO Financial Corp.	Bank Holding Company	1.03	0.09%
KeyCorp	Bank Holding Company	1.00	0.09%
M&T Bank Corporation	Bank Holding Company	0.83	0.07%
Huntington Bancshares Incorporated	Bank Holding Company	0.78	0.07%
Commerce Bancshares, Inc.	Bank Holding Company	0.59	0.05%
Coastal Community Bank	Commercial Bank	0.53	0.05%
1st Financial Bank USA	Commercial Bank	0.39	0.03%
Santander Holdings USA, Inc.	Bank Holding Company	0.35	0.03%
First BanCorp.	Bank Holding Company	0.33	0.03%
First Citizens BancShares, Inc.	Bank Holding Company	0.33	0.03%
SoFi Technologies, Inc.	Bank Holding Company	0.32	0.03%
RBC US Group Holdings LLC	Bank Holding Company	0.31	0.03%
UBS Americas Holding LLC	Bank Holding Company	0.29	0.03%
First National Bank	Commercial Bank	0.23	0.02%
First Hawaiian, Inc.	Bank Holding Company	0.22	0.02%
TCM Bank, National Association	Commercial Bank	0.21	0.02%
Lead Bank	Commercial Bank	0.20	0.02%
Arvest Bank Group, Inc.	Bank Holding Company	0.20	0.02%
HSBC North America Holdings Inc.	Bank Holding Company	0.20	0.02%
Synovus Financial Corp.	Bank Holding Company	0.19	0.02%
UMB Financial Corporation	Bank Holding Company	0.19	0.02%
Simmons First National Corporation	Bank Holding Company	0.16	0.01%
Zions Bancorporation, National Association	Commercial Bank	0.16	0.01%
Associated Banc-Corp	Bank Holding Company	0.14	0.01%

Table 1: Market concentration of the credit card industry by tracking the top 50 depository institution issuers in 2024 (Calem and Gross).

In addition to price competition, card issuers compete through non-price strategies, most notably rewards programs. Data from IBISWorld estimates the value of the credit card rewards program industry to be $952.0 million in 2025. Between 2019 and 2024, expenditures on rewards programs by issuers have increased at a compound annual growth rate (CAGR) of 2.4%, and is expected to increase by 1.1% in 2025 alone (IBISWorld). The scale and growth of credit card rewards program expenditures suggest that competition in the credit card market operates through non-price means, potentially reducing the role of APR adjustments as the primary instrument of competition.

Discussion of Results

Credit card APRs increased strongly with the Fed's tightening cycles, increasing 2.51% during the November 2015–November 2018 tightening and rising 6.63% during the March 2022–July 2023 cycle. The spread between APRs and the EFFR remained relatively stable throughout these periods, suggesting that increases in benchmark rates were largely transmitted to consumer lending rates unimpeded.

During monetary easing cycles, credit card APRs fell by 1.47% between August 2007–November 2008, 0.58% from August 2019–May 2020 cuts, and 0.79% between August 2024–November 2025 cycle. During these periods, the spread between APRs and the EFFR widened considerably, indicating that declines in benchmark rate were not passed through to consumers.

During periods of EFFR increases, credit card APRs are rising at an average rate of 4.57% over the course of that cycle. By contrast, over the length of EFFR cuts, credit card APRs are falling at an average rate of 0.95%. The data indicate that pass-through downwards is asymmetrical relative to upwards: increasing 4.57% following EFFR hikes and decreasing 0.95% following declines in the EFFR.

The trend depicted in Figure 2 buttresses this hypothesis. In this context, a stable spread indicates relatively proportional adjustment to rate changes, however, a widening spread is characterized by APRs adjusting more slowly than changes in monetary policy rates. The difference between the average credit card APR rose rapidly throughout and following the conclusion of each easing period, whereas the spread remained roughly constant during the tightening periods.

Market structure amplifies this effect. With a four-firm concentration ratio (CR4) of 53.03%, the U.S. credit card market exhibits oligopolistic characteristics that theory predicts will lead to price stickiness (Neumark and Sharpe). Thus competition will shift to non-price forms, relying on rewards programs, rather than APR reductions, to be the primary differentiator (Siddharthan). High market concentration is consistent with reduced competitive pressure to pass lower benchmark rates onto consumers, reinforcing the observed asymmetry (Hannan and Berger). IBISWorld data supports this, highlighting that expenditures on rewards programs by retailers have increased at a compound annual growth rate (CAGR) of 2.4%, and 2025

expenditures total $952 million as of January 2025 (IBISWorld), highlighting a clear shift toward non-price competition.

APRs competition affects marginal, often riskier borrowers, while rewards competition targets high-spending, low-default, and inelastic borrowers. Economic theory predicts that firms prefer to compete in markets where demand is less price sensitive, profits are higher and risk is lower.

APR reductions apply to all revolving balances, thus reducing revenue across all balances, unlike rewards, which can selectively target specific customer segments, rendering APR reductions a blunt form of competition that firms seek to avoid. Rewards programs better match these criteria, therefore, firms will seek to substitute non-price competition for price competition whenever possible. Thus, in concentrated markets, rewards programs become the focal point of competition, rather than APRs. As issuers shift towards rewards-based competition during easing cycles, competitive pressure is redirected away from APR reductions, allowing the APR-EFFR spread to widen despite falling benchmark rates.

Admittedly, some industry groups, such as the Bank Policy Institute, argue that the credit card industry is highly competitive, citing the presence of over 4,000 issuers and the rise of fintech alternatives. However, while the total number of issuers is high, the CR4 ratio of 53.03% indicates that a small group of dominant issuers hold significant pricing power over the majority of the market. For the average consumer, competition from small banks does not result in lower APRs, as evidenced by the widening APR-EFFR spread observed.

Limitations

The analysis methods used in this paper are subject to several limitations. First, credit card APRs are reported at discrete intervals and do not perfectly align with the beginning and ending of each Federal Reserve policy regime. As a result, data is taken from the closest available month, to each monetary phase. While this introduces minor timing imprecision, the magnitude and persistence of the APR-EFFR spread across several monetary cycles indicate that results are not merely driven by timing alone.

Second, the analysis relies on aggregate APR data, which is unable to account for issuer-level and consumer-level data. Therefore, this analysis is unable to account for individual-level factors that may influence pass-through rates. However, the use of aggregate data is appropriate for this analysis, as the goal is to broad transmission patterns, rather than individual outcomes.

Third, this analysis lacks any formal econometric analysis and instead employs a descriptive time series analysis. Consequently, any conclusions should be interpreted as evidence of observed correlations and structural patterns, rather than causal. Nonetheless, the consistency of asymmetric pass-through across multiple monetary cycles and its alignment with previous literature, support the interpretation that market structure plays a meaningful role.

Further research applying econometric methods to more precise issuer-level data could more accurately examine the extent to which market structure inhibits interest rate pass-through in the U.S.

Conclusion

Taken together, the evidence indicates that credit card APRs experience asymmetrical pass-through: APRs increased 4.57% on average following EFFR hikes but only decreased 0.95% on average following declines. This pattern is consistent with an oligopolistic credit card industry, where prices are sticky and competition is primarily driven through rewards programs and other non-price strategies. Consequently, monetary easing imperfectly transmits to consumer borrowing costs, implying that benchmark rate cuts provide less relief to borrowers than hikes impose during tightening regimes. These findings illustrate the limitations of monetary policy in concentrated consumer markets, suggesting that policymakers should account for structural factors when evaluating the effectiveness of rate cuts.

Works Cited

Agarwal, Sumit, et al. “Do Banks Pass through Credit Expansions? The Marginal Profitability of Consumer Lending during the Great Recession.” SSRN Electronic Journal, 2015, https://doi.org/10.2139/ssrn.2676014.

Bernanke, Ben S., and Mark Gertler. “Inside the Black Box: The Credit Channel of Monetary Policy Transmission.” Journal of Economic Perspectives, vol. 9, no. 4, Nov. 1995, pp. 27–48, www.aeaweb.org/articles?id=10.1257/jep.9.4.27. Accessed 13 Jan. 2026.

Board of Governors of the Federal Reserve System (US). “Commercial Bank Interest Rate on Credit Card Plans, All Accounts.” FRED, Federal Reserve Bank of St. Louis, 1 Nov. 1994, fred.stlouisfed.org/series/TERMCBCCALLNS. Accessed 3 Jan. 2026.

Calem, Paul, and Benjamin Gross. “The Credit Card Market Is Not Even close to Being Overly Concentrated - Bank Policy Institute.” Bank Policy Institute, 18 Apr. 2024, bpi.com/the-credit-card-market-is-not-even-close-to-being-overly-concentrated/. Accessed 10 Jan. 2026.

Federal Reserve. “The Fed - Consumer Credit - G.19.” Federalreserve.gov, 2019, www.federalreserve.gov/releases/g19/current/. Accessed 3 Jan. 2026.

Federal Reserve Bank of New York. “Secured Overnight Financing Rate Data - Federal Reserve Bank of New York.” www.newyorkfed.org, 2024, www.newyorkfed.org/markets/reference-rates/sofr. Accessed 3 Jan. 2026.

Gropp, Reint, et al. “The Dynamics of Bank Spreads and Financial Structure.” The Quarterly Journal of Finance, vol. 04, no. 04, World Scientific, 2014, pp. 1450014–14, https://doi.org/10.1142/s2010139214500141. Accessed 13 Jan. 2026.

Hannan, Timothy H., and Allen Berger. “The Rigidity of Prices: Evidence from the Banking Industry.” Journal of Banking and Finance, vol. 15, no. 1, Feb. 1991, pp. 133–49, https://doi.org/10.1016/0378-4266(91)90042-k. Accessed 13 Jan. 2026.

IBISWorld. “Card Reward Program Services in the US - Market Research Report (2015-2030).” Ibisworld.com, 2015, www.ibisworld.com/united-states/industry/card-reward-program-services/5842/. Accessed 8 Jan. 2026.

Irby, LaToya. “How Do Fed Interest Rate Cuts Impact Credit Cards?” Experian.com, Experian, 30 Jan. 2024, www.experian.com/blogs/ask-experian/how-do-fed-rate-cuts-impact-credit-cards/. Accessed 13 Jan. 2026.

Karagiannis, Stelios, et al. “Interest Rate Pass-through in Europe and the US: Monetary Policy after the Financial Crisis.” Journal of Policy Modeling, vol. 32, no. 3, May 2010, pp. 323–38, https://doi.org/10.1016/j.jpolmod.2010.02.006. Accessed 13 Jan. 207AD.

Kwapil, Claudia, and Johann Scharler. “Interest Rate Pass-Through, Monetary Policy Rules and Macroeconomic Stability.” Journal of International Money and Finance, vol. 29, no. 2, Mar. 2010, pp. 236–51, https://doi.org/10.1016/j.jimonfin.2009.06.010. Accessed 8 Jan. 2026.

McCann, Adam. “Credit Card Market Share by Issuer (2022).” WalletHub, 9 Apr. 2025, wallethub.com/edu/cc/market-share-by-credit-card-issuer/25530. Accessed 4 Jan. 2026.

Mishkin, Frederic. The Channels of Monetary Transmission: Lessons for Monetary Policy. Feb. 1996, www.nber.org/system/files/working_papers/w5464/w5464.pdf. Accessed 14 Jan. 2026.

Neumark, David, and Steven Sharpe. “Market Structure and the Nature of Price Rigidity: Evidence from the Market for Consumer Deposits.” The Quarterly Journal of Economics, 1 May 1992, academic.oup.com/qje/article-abstract/107/2/657/1838340. Accessed 10 Jan. 2026.

“Regulating Consumer Financial Products: Evidence from Credit Cards *.” The Quarterly Journal of Economics, vol. 130, no. 1, Nov. 2014, pp. 111–64, https://doi.org/10.1093/qje/qju037.

Rossman, Ted. “Current Credit Card Interest Rates.” Bankrate, Bankrate.com, 4 Oct. 2024, www.bankrate.com/credit-cards/advice/current-interest-rates/. Accessed 10 Jan. 2026.

S, Siddharthan N. “Industrial Structure, NonPrice Competition and Industrial Development.” Economic and Political Weekly, vol. 19, no. 31/33, Economic and Political Weekly, 1984, pp. 1307–10, https://doi.org/10.2307/4373484. JSTOR. Accessed 10 Jan. 2026.

Scholnick, Barry. “Interest Rate Asymmetries in Long-Term Loan and Deposit Markets.” Journal of Financial Services Research, vol. 16, no. 1, Springer Science+Business Media, Jan. 1999, pp. 5–26, https://doi.org/10.1023/a:1008107030893.

The Consumer Credit Card Market Report. Consumer Financial Protection Bureau, 25 Oct. 2023, files.consumerfinance.gov/f/documents/cfpb_consumer-credit-card-market-report_2023.pdf. Accessed 16 Jan. 2026.

“U.S. Credit Card Industry: Competitive Developments Need to Be Closely Monitored.” Www.govinfo.gov, 28 Apr. 1994, www.govinfo.gov/content/pkg/GAOREPORTS-GGD-94-23/html/GAOREPORTS-GGD-94-23.htm. Accessed 13 Jan. 2026.

The Evolution of Sports Contracts and the Role of Agents in Negotiations
By Ajay D'Souza

Abstract

Sports contracts have shifted dramatically over the past century, evolving from basic one-page deals into detailed legal agreements that shape athlete mobility, compensation, and control over career decisions. This paper traces those changes across major American and international leagues, emphasizing the growing role of player unions, the introduction of free agency, and the rise of professional sports agents. It examines how key institutional rules-from MLB's reserve clause to the NBA's salary cap and European football's transfer systems-have influenced negotiating power. The paper also explores how artificial intelligence is beginning to impact contract valuation and labor negotiations. By combining historical trends with current developments, this study offers a clear overview of how contracts have changed and where they're likely headed.

Introduction

Modern sports contracts are among the most complex legal tools in entertainment and business. But their earliest forms were far simpler, offering athletes little say over their salaries, team assignments, or job security. As sports became more commercial and labor protections grew throughout the 20th century, contracts evolved to reflect new ideas about fairness, competition, and professional support (Smith). Agents and legal advisors emerged as key figures in this environment, helping athletes navigate increasingly complex rules. More recently, new tools like artificial intelligence have begun shaping how teams and players assess value and risk. This paper tracks those shifts-how contracts were built, why they changed, and what they reveal about the balance of power between players and organizations.

Early Sports Contracts (1900-1960)

In the early 20th century, sports contracts heavily favored teams. Players typically signed one-year deals with little room to negotiate. The most significant restriction was MLB's reserve clause, which allowed teams to hold players indefinitely, even after contracts expired (Flood). It was pitched as a way to maintain parity, but it effectively stripped players of leverage. Without unions or legal protections, those who pushed back risked being blacklisted (Flood; Grange). These contracts mirrored broader labor norms of the era, where workers had few rights and employers held most of the power.

The Rise of Free Agency and Player Unions (1960s-1980s)

The 1960s and '70s saw major change. Broader labor movements began influencing sports, and players started organizing. A key turning point came with MLB's Curt Flood, who refused a trade and challenged the reserve clause in court (Flood). Though he lost, his actions sparked momentum. In 1975, arbitrator Peter Seitz effectively ended the reserve system, paving

the way for free agency (Flood). At the same time, player unions in the NFL, NBA, and NHL negotiated collective bargaining agreements (CBAs) that improved pay, introduced grievance systems, and created free market conditions (Smith). Teams had to compete for talent, and agents became essential in helping players navigate salary caps, arbitration rules, and long-term planning.

Professionalization of Sports Agents (1980s-2000s)

By the 1980s, sports agents had moved from informal advisors to powerful negotiators. Figures like Scott Boras reshaped the industry, using legal strategy, statistical analysis, and long-term planning to land record-setting contracts (Boras). Agents no longer focused solely on salaries - they also advised clients on endorsements, branding, and career direction (Boras). As salaries grew and league rules became more complex, agents had to understand not just the law but also economics and cap mechanics. They became key intermediaries, balancing player priorities with financial and regulatory realities.

Case Studies

- LeBron James: By signing short-term deals with player options, James kept flexibility front and center. With Rich Paul's guidance, he maintained control over his trajectory while maximizing leverage (Smith).

- Patrick Mahomes: His 10-year deal with the Kansas City Chiefs included cap-friendly structures and incentives. It shows how long-term contracts can still offer strategic control (Rosenthal).

- Lionel Messi: His departure from FC Barcelona in 2021 highlighted how European football regulations - like financial controls and salary limits - can override even high-profile deals (Smith).

- Bryce Harper: Harper's 13-year deal with the Phillies emphasized long-term security over yearly flexibility, showing how personal values shape contract decisions (Fitzpatrick).

Comparing Leagues: Structure and Power

Contract terms and player leverage vary widely between leagues. The NFL's non-guaranteed deals make signing bonuses and early guarantees crucial (Smith). In contrast, NBA players benefit from more guaranteed money and contract flexibility. In European football, agents must navigate cross-border labor laws, buyouts, and transfer fees - making negotiations even more complex (Smith). These differences show that athlete power depends as much on league structure and legal frameworks as it does on individual performance.

AI's Growing Role in Contracts

Artificial intelligence is starting to influence contract negotiations, particularly in performance analysis, injury forecasting, and market valuation (Wilson). These tools can streamline decisions and surface hidden trends, but they also raise concerns - especially around data privacy, algorithmic bias, and fairness (Wilson). As AI becomes more embedded in sports, its impact on contracts and labor negotiations will expand, shaping both how deals are made and how disputes are resolved. Future research should focus on how AI-driven metrics influence CBAs. Who controls the data? How transparent are the models? These questions will matter as leagues and unions adapt to new technologies. Getting the balance right will be essential to avoid reinforcing existing power imbalances.

Conclusion

The history of sports contracts reflects broader shifts in labor, economics, and technology. What started as one-sided agreements has become a complex field shaped by law, negotiation, and analytics. Agents remain critical in helping players navigate this world, and new technologies are starting to redefine how value is measured. As professional sports continue to evolve, contracts will remain a key site where power, money, and strategy intersect.

Works Cited

Boras, Scott. The Business of Player Representation. Sports Legal Press, 2019.

Flood, Curt. The Way It Is. Trident Press, 1971.

Fitzpatrick, Frank. “Bryce Harper’s Contract: A Game-Changer in MLB Free Agency.” Sports Business Journal, 2019.

Grange, Red. Red Grange Remembers. Simon & Schuster, 1953.

Rosenthal, Ken. “Patrick Mahomes’ Historic Contract and Its Impact.” The Athletic, 2020.

Smith, Sam. The Business of Basketball Contracts. Harvard University Press, 2021.

Wilson, James. “AI and the Future of Sports Contract Negotiation.” Journal of Sports Law & Business, 2023.

The Sudanese Civil War: Physical, Psychological, and Global Outcomes By Lilla Prosko

Abstract

The current Sudanese civil war that began in 2023 has severe and lasting impacts on civilians, and is seeing a lack of aid and international awareness. Its death toll rises every day due to starvation, disease, bombing, and more, reaching far over 150,000 people while at least 11 million people have been displaced already. It is essential to consider what the physical and psychological effects on people impacted during and after the Sudanese civil war are, as well as the role that greater awareness plays in ameliorating this suffering. The current civil war is Sudan's third in history, and began in 2023. The Rapid Support Forces was formed due to government dispute, and opened fire on the Sudanese Army. Reporters have been able to enter the country and document the horrors civilians are facing, providing access to many interviews and images that will be referenced as firsthand accounts. However, the difference in popularity and awareness of this war compared to other current conflicts is drastic, and is developed through the use of data and Google Trends charts. Civilians see the physical impacts, as they are victims of war crimes such as hospital bombings, rape, and forced labor. Those who stay in Sudan face challenges such as destruction of infrastructure and widespread malnutrition, especially among young children. Every Sudanese civilian has been traumatized, facing dehumanization and loss of hope for their future. The lack of awareness of the Sudanese civil war correlates to the insufficient humanitarian aid they have been barely receiving, which raises questions of selective aid distribution and American exceptionalism. Increased awareness is essential to help the civilian situation, and many international organizations are working to provide supplies. The humanity of these people is an urgent matter, and it is not too late to improve the current aid that is too little.

Introduction

"People need to know what's happening"

- Anonymous Sudanese civilian who was forced to spy for the RSF in exchange for her children's lives

An American news broadcaster announces the week's highlights in a loud, affirmative voice. He reports on the weather and which baseball team has the most wins. The broadcast concludes with a segment on another online trend gone awry. Once again, the news broadcast conveniently says nothing about the millions of people throughout the world who contend with war and its extensive consequences.

Sudan is one example of a long running, and indisputably devastating war. In the rare circumstance that the Sudanese Civil War is brought to the attention of Americans, most viewers turn away from the opportunity to learn with the click of a button from a hand held remote. The large, flat screen will remain silent, collecting dust, just like the abandoned household items left

behind by those fleeing their homes in Sudan. In Sudan, dust is everywhere. It blows across streets when shells from aerial bombing land in towns across the country. It is packed into bags and stacked in front of hospital rooms where windows used to be. People cough up dust as they cross miles of land, seeking a safe haven, only stopping to bury their loved ones throughout the journey. These civilians have lost the luxury of turning off a television, long ago having lost the luxury of ignorance.

While most people who are thousands of miles away understandably fail to comprehend the vast, aggressive force that is war, the effects of the Sudanese Civil War are profound, far more than dust storms and abandoned towns. Civilians experiencing war have important stories to tell, if only people choose to listen. So, what are the physical and psychological effects on people impacted during and after the Sudanese civil war? Moreover, what role does greater awareness play in ameliorating this suffering?

Having been fortunate enough to travel outside the United States, I have seen the contrast in countries across the world. While there are countries similar to the US that I have visited, such as Italy or Australia, I have clear memories of visiting Egypt, my first time on the continent of Africa. I observed poverty in a variety of ways, but the tragedy of the situation was the same. People offered items in exchange for whatever they could get, and single motorcycles were shared among large families. Many people's only earnings came from tourists. I left the country with more questions than answers about the balance of resources in the world and global cooperation. Egypt's poverty demonstrated the unequal distribution of international aid. If there was a system of aid I did not know about, why was I not seeing the effects on civilians? I had visited many countries that had resources to help, including my own country, so I failed to understand the lack of aid. However, there may be a much simpler explanation. I am one of the relatively few Americans who have the opportunity to visit underdeveloped countries and encounter injustice in the world. If people do not know about something, they will not act on it. Since many Americans lack the freedom and resources to personally visit countries outside the United States, it is essential to spread knowledge. Similarly, poverty in the United States is often overlooked and ignored; the imbalance of wealth is present everywhere and, more importantly, the actions people are not taking to help others. After reading about injustice happening in a different African country, I understood the importance of spreading knowledge across the world in order to assist others, even if they are far away. I aim to raise awareness about the suffering of Sudanese civilians in hopes of protecting their human rights.

Sudan has a history of civil wars, and is currently fighting its third. The first civil war began in 1955 due to a dispute between Northern and Southern regions of the country. Ending in 1972, the death toll amounted to 500,000 people. However, the conflict was not over. The second civil war began in 1983 and resulted in a death toll of 2 million at its end in 2005. At this point, Sudan has officially been split into two countries: Sudan and South Sudan. After the split of the nation into two sovereign countries, it comes as a surprise that Sudan is now in its third civil war. However, Sudan is very difficult to govern due to the ethnic differences and diverse groups in the country. This wide range of people has complicated politics, as all groups want to be represented

in the government. When people are unsatisfied with the decisions of their government, they can branch off in extreme situations as is the case with Sudan's third civil war. Additionally, Sudan's failure to rebuild the country after war leaves it vulnerable to future conflict. Sudan's government has taken actions after the first two civil wars to unify its country, but little success has been observed. After a long period of trying to rebuild its region and government, there were attempts at democracy. The people of Sudan were frustrated by the lack of representation of the many ethnic groups in society. Through their desire to be represented and accepted as a diverse community, the citizens pushed for this new democratic system. Unfortunately, these demands were not met and the move towards democracy was vanquished. Two powerful Sudanese military generals had been having a personal dispute that led to one of them breaking off from the government and creating a militia known as the Rapid Support Forces (RSF). In 2023, the RSF opened fire, marking the start of another civil war for Sudan. The Sudanese Army has continued to fight back against the RSF and finally shifted to an offensive side. Since then, the war has continued to rage as civilians are being killed and displaced. As many attempt to flee their towns and country, they face significant barriers. Those that stay are packed into schools turned into living quarters or hospitals. They live off aid that has been offered from volunteers and international agencies, which is barely enough to get civilians to the next day. As these people suffer through the tragedies of war, they lose hope and struggle to imagine a bright future.

Current Situation

While violence and destruction continues, the awareness of this atrocity has seen no progress. Many Sudanese civilians must eat leaves to survive while most other countries remain in blissful ignorance. Smaller scale warfare seems to be classified as less significant to other parts of the world, especially if they do not benefit. For example, the US did not fail to cover news on the Russia-Ukraine conflict, specifically when a drop in gas prices was observed. Even the conflict in Gaza is well-known in households across the United StatesAmerica. Some countries such as Iran, Saudi Arabia, and the United States send aid through weapons, but fail to get more involved. There has been no interference or reports of soldiers being sent to help. As well as soldiers, very little health and welfare resources have made their way to the Sudanese civilians. At this time, there is a desperate need for anything other than weapons; yet weapons are all that is being sent to Sudan from international powers. Civilians have experienced more than two months at a time without any sign of international aid. The war still rages on, and people around the world are still oblivious to this situation. If people have no knowledge, they are unable to help. In times of need, aid and awareness are especially important, and Sudan's third civil war is no exception.

However, in past times, historical responses differed. Darfur, a region in Sudan, became a household name in the early 2000s due to a catastrophic genocide. It was categorized as an ethnic cleansing, occurring between 2003-2008. At the time it was common to hear about Darfur as an example of genocide.Yet few people were informed on Sudan and its three civil wars. Similar to its past, Darfur now experiences turmoil in the Third Sudanese Civil War due to its

importance to the country. The region, especially its city, El-Fasher, was controlled by the Sudanese Army and was one of the most essential territories to the army. However, the city was officially seized by the RSF on October 26, 2025. While the RSF had been cutting off residents from crucial supplies such as food and medicine for 18 months, this siege of the city has already displaced an additional 470,000 people. The future of these civilians and the power of the Sudanese Army is uncertain, and will not improve unless Sudan sees international aid, and soon.

People in a developed country tend to take basic necessities for granted. Even the people of Sudan did not understand the horrors they were about to face when war began. Access to these basic necessities has been stripped of many civilians and they face health problems daily. This includes malnutrition, especially for many young children and babies. With fragile bodies, they can barely survive the week. Sudan's Civil War is putting the country on the brink of famine, proven by the 3.5 million children under the age of five that suffer from malnutrition (BBC News Africa). In the villages that have not been deserted, people are living off rations and supplies donated by volunteers in the community. Without any outside aid, Sudanese civilians are left to help each other. The RSF has only made matters worse, capturing areas of high agricultural density, known as the "bread basket" of the country. They have cut off access to these goods to civilians, leaving people to turn to harsh methods to survive. These include eating things such as leaves and soil as it has become their only source of 'nutrition' during the war. After many years of war, some citizens have become accustomed to eating leaves. A Sudanese citizen tells a reporter from the New York Times that "no one has food. We are all eating bush leaves" (Casey). Everyone in the village contributes to the process, some breaking off branches while others collect them in bags to be boiled for consumption later. This is their only source of nutrition, and has been for a while. The citizen expresses how this has become the situation for almost all the locals. While they flee from the crossfire and tend to their injured companions, they are also struggling with food. There is little left for these people, and there is no telling what citizens will resort to when even the leaves become scarce.

“Leaves gathered for food in a camp for displaced Sudanese” (New York Times) https://www.nytimes.com/2024/08/08/magazine/sudan-nuba-war.html?smid=nytcore-ios-share&referringSource=articleShare

After running away from their homes due to the violence, many families have been forced to find safety elsewhere. This includes schools transformed into housing, often leaving one room per family. The cramped areas and decrepit conditions are grounds for sickness and disease, leading to deadly cholera outbreaks in many of these ‘villages.’ Children are no longer receiving an education and have not received an education since the beginning of the war—almost three years. Due to many families living together, there is a limited amount of resources. Each person has one meal a day, if they are lucky. The communities also live in constant fear of attacks, even though they have already fled their previous homes. It has become dangerous to stay in one place for a longer period of time due to the nature of attacks and RSF bombing. People are also losing access to necessary daily medication. One woman living in a schoolroom with her family expresses that she has a condition requiring medication, but she has not received this since the beginning of the war and her health is declining. Another man, who was captured by the RSF, tells reporters from the *Wall Street Journal* that he is diabetic and the RSF would not even give him water, much less his necessary medical supplies. He tells the reporter “I thought I was going to die . . . I was drinking my own urine for three days” (The Wall Street Journal 11:32-11:50). This man was held by the RSF for twenty days because they believed he had connections with Sudanese Army soldiers due to a photo on his phone. He was not even a soldier, yet was subjected to unimaginable horrors that will forever imprint on his mind, however long he has left. The lifestyle of so many citizens has been shaken up by the war; they lose hope and family members every day.

In the first year of war, 10,000 people were killed. As of 2025, the death toll has progressed to over 150,000. The cause of death can be attributed to many factors such as starvation, exhaustion, and disease, but aerial bombing has become a much more common method of violence in the Sudanese Civil War. So common that it is used by *both* the RSF and the Sudanese Army. The effect of the violence has emotionally damaged civilians, as many express they are losing motivation for survival and claim they would rather die where they are than have to run away or be displaced again. Violence does not stop at killing, however. There are many war crimes that the RSF is being accused of, including rape, attacking hospitals, and capturing civilians. Those civilians held captive by the RSF are forced into working for them by fighting or being spies. These people have experienced blackmail, specifically in one case where a woman was forced to be a spy for the RSF or they threatened to kill her children. After escaping their grasp, she exclaims that “people need to know what’s happening” (PBS News Hour). As just one of the many instances where people are forced into war, the woman voices the concerns of her fellow citizens. There is little to no help being provided to Sudan, and she considers herself one of the few who managed to escape the RSF alive.

The RSF has also been extending its violence outside of Sudan. They are kidnapping mercenaries from South Sudan and forcing them to fight without any form of compensation. Mercenaries from Chad have also expressed how they were tricked into working for the RSF by the promise of money in exchange for their services. When they arrived, the RSF made it clear the arrangement would not be happening and the mercenaries had no way to leave. Women have been sexually assaulted and some accusations against the RSF claim they have been trading or selling women. A reporter from BBC News visited the front lines of the war, escorted and given permission by the Sudanese Army. There, she interviewed many Sudanese civilians whose identity will remain confidential. However, their stories are very impactful. One woman recalls how the RSF broke into her house and demanded to rape her two daughters, the youngest daughter being only ten years old. Her remarkable bravery saved her daughters, as she told the RSF soldiers, "if you want to rape anyone it has to be me" (BBC News 4:00-4:03). Unfortunately this case is seen all throughout the war, and is not a unique experience among Sudanese women. The reporter then meets a group of women in the Omdurman City who are leaving RSF territory. They too have experienced similar events, as one woman exclaims "there are women here who've been raped, but they don't talk about it. They hide it. What difference would it make if they speak?" (BBC News 1:12-1:19). People have lost hope for change, and fear their country will not recover from this war. They already carry wounds that cannot be healed, ones which have damaged their soul and not their skin.

Apart from the war crimes committed to individuals, bombing hospitals is also a war crime. Hospitals have been trying to save young children facing severe malnourishment but lack the resources. They also see many civilians who have been hit in the crossfire of the war, as one man recalls how he became paralyzed by a shell. He tells a reporter from BBC News that "the bomb hit the wall in front of me. My friend screamed for help. There was blood coming out of my nose and mouth" (BBC News 3:13-3:23). He is still a young boy around age 15, but has been permanently handicapped by the violence. His mother expresses her sorrow about how his quality of life is changed forever. The young man, still in the hospital, has a chance of dying. This is just one of the accounts where many people try to seek help and safety in hospitals, but fear they may be exposed to more violence due to the RSF. The RSF has tried to defend themselves by claiming they believed the hospitals were housing soldiers, but this defense was quickly proved weak. Even so, places with many civilians are still being hit. Hospital nurses struggle to save lives while keeping the area secure. This includes packing windows with bags of sand in an attempt to stop shells from exploding inside the buildings.

A Sudanese hospital packed with civilians who are victims of bombings (Wall Street Journal) https://www.youtube.com/watch?v=LtLszFIFPUo

Even though it is considered a war crime, many of these attacks target civilians. Hospitals are not filled with soldiers, but civilians. A Wall Street Journal reporter visits another hospital in Khartoum, where a staff member tells him about the horrors they encounter. After walking through the building, the staff member sheds light on the inhumanity of war, saying that the RSF is "just killing people for nothing. You saw the victims yourself. They are not soldiers. They are not armed. They are kids." (The Wall Street Journal 9:14-9:20). And yet, they are losing siblings, parents, and even their own lives to the effects of violence. This greatly decreases the standard of life in war and affects the Sudanese people. Sudan has become the ground for street-to-street urban warfare, leading to the displacement of millions of civilians. Specifically, at least 11 million people have been displaced so far. This number continues to climb as the war rages on. Over 40 villages were burned in just two months of war, attributing to the high displacement number since countless people do not have a home anymore. Their homes and communities have been so damaged that people are no longer able to properly bury their loved ones. They have resorted to quickly burying family members in the dirt whenever they can. One man, the sole person who remained in his village, shows the burial grounds to a reporter from BBC News Africa. He explains this makeshift burial ground contains most of the civilians in his village, and that "they all died from Rapid Support Forces bombs . . . the Rapid Support Forces forbade their burial in the main cemetery" (BBC News Africa 5:02-5:13). He shows his friends have become mounds in the dirt marked by numbers, counting the amount of bodies in the area. They no longer have names on their gravemarkers due to the rush of escaping the crossfire. The war in Sudan is stripping people of their identity, similar to events observed in the past. People affected by the Holocaust, a catastrophic ethnic cleansing, were tattooed with codes containing multiple numbers. This was a harsh dehumanization process, and the parallels between names being replaced by numbers in Sudanese gravesites sheds light on the true horrors of the civil war. Such

similarity to one of history's greatest tragedies displays how people are affected by this war; at the end of it, they will not even be remembered correctly. While the man may remember the names of his friends buried in his village, these people have lost their identity as individual people after their burial and are now left in history as grave #147, #148, #149, and so on.

Graves marked with numbers 101 and 139 in a makeshift graveyard on the outskirts of Omdurman, which has been almost completely deserted (BBC News Africa 4:38) https://youtu.be/KIDMsalYHG8?si=nSVv4VtGEU5XEl0n

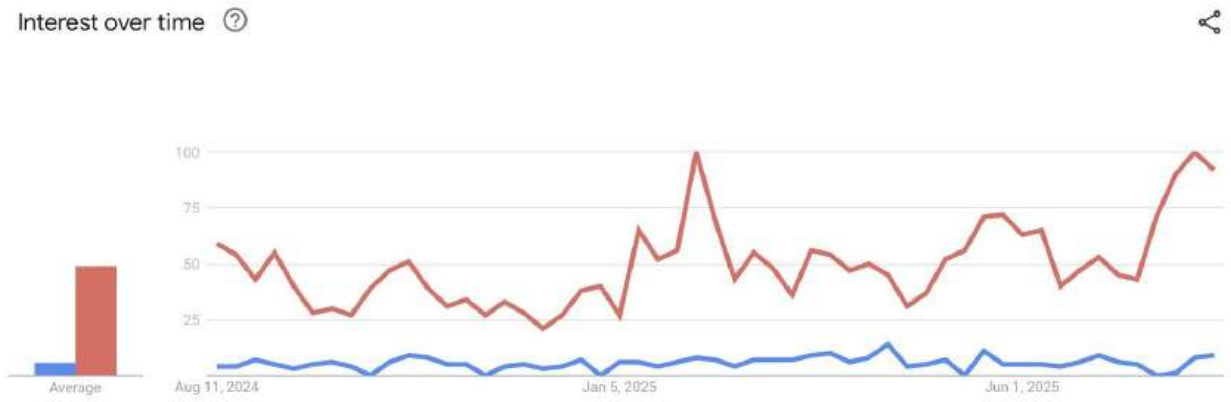
International Involvement

The effects of the Sudanese Civil War are grand, and yet they seem to go unnoticed by most of the world. However, some countries have taken notice. Countries such as the US, Russia, Iran, and the UAE have sent resources. The US, Russia, and Iran have sided with the Sudanese Army by sending weapons, specifically from Russia and Iran. There is not much information pertaining to how effective and at what scale their weapons seem to be aiding the fighting. There are reports of old weapons from previous Sudanese Civil Wars being found and modified by the RSF to be used in the fighting. The US supports the Sudanese Army, but has yet to get directly involved or send soldiers. On the other hand, the UAE has been accused of supplying weapons to the RSF. A suspected reason for this occurrence is that the UAE wants control over Sudan's ports and mines, which the RSF is willing to comply with should they win the war. Especially Port Sudan, which is an essential area for exportation and importation, as well as a key to receiving international supplies. Although there is this attempt for international reach to Sudan, it becomes difficult due to legal restrictions and violence throughout the country and in certain essential areas. Sudanese citizens claim they have gone as long as two months without seeing any form of international support. During these times, there is a desperate need for humanitarian aid from outside powers. This need remains unmet.

In the first year of war, the US had tried to mediate peace talks and find resolution, yet neither opposing side managed to show up to the attempted peace meeting. Both sides seem to have strong feelings about a resolution being impossible peacefully. At this point, there may not

be a peaceful way of achieving an end to the war; it will result in the RSF or Sudanese Army being completely defeated physically, and only time will tell who. As conflict first broke out, the Vice President of Sudan, Malik Agar, spoke with PBS News Hour. Agar demonstrates the belief that peaceful resolution is near to impossible, stating that "any leader in Sudan who can go into negotiations with the RSF is committing a political suicide" (PBS News Hour 2:34-2:40). There is fear among civilians that, in terms of help, it is "too little, too late" (Al Jazeera English 11:26-11:28).

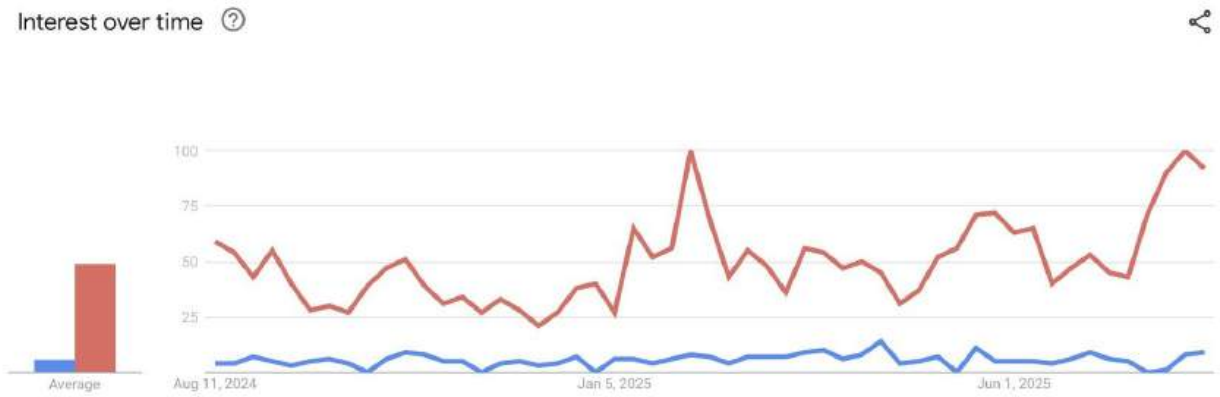
It is a phenomenon that the Sudanese Civil War still rages on even after receiving aid from some powerful countries. However, when looked at from a closer view, it becomes apparent why. While there is aid being provided by other countries, it is not as much as most other conflicts. This civil war is also less known among people worldwide, which may link to the amount of aid given internationally. For example, the conflicts in Ukraine and Gaza have both received more international aid and media coverage than Sudan. Additionally, people's curiosity is spiked more by those other conflicts than by Sudan itself. When the Sudanese fighting first broke out in April 2023, six months before Gaza's conflict began, it was one of the very few occurrences of 'Sudan' being searched more by people than 'Gaza'. Even so, the average search on Gaza immediately spiked and reporting on Sudan was again buried. Since then, Gaza has had far more coverage. Sudan has never reached the height of recognition that Gaza has in the United States and is not looking like it will in the near future.



Google Trends chart of the frequency of search terms 'Sudan' (blue) and 'Gaza' (red) in the US from the past 12 months
https://trends.google.com/trends/explore?date=all_2008,all_2008&geo=US,US&gprop=news&q=Sudan,Gaza&hl=en

Similarly, Ukraine sees more online popularity than Sudan. In relation to Ukraine's online presence in the US, Sudan has almost none. When the terms 'Sudan' and 'Ukraine' are compared side by side, Ukraine dwarfs Sudan throughout the past year. The charts provided by Google Trends exposes just how little the community in the US is concerned with Sudan. The amount of

media coverage can also be linked to the impact on other countries, as the Russia-Ukraine conflict became very commonly discussed in America when it affected gas prices around the country. Therefore, more people are searching the web for topics that directly affect them, and remain careless to anything else. People's interest level also dies off after a certain period of time. Spikes in search trends are present due to the initial interest and curiosity of many people. After their questions have been answered or their worries reassured, most US citizens resort back to their regular routine because they are not personally seeing the effects of the tragedies in war.



Google Trends chart of the frequency of search terms 'Sudan' (blue) and 'Ukraine' (red) in the US from the past 12 months
https://trends.google.com/trends/explore?date=all_2008&geo=US&gprop=news&q=Sudan,Ukraine&hl=en

While there is still news coverage, less attention in day-to-day conversation becomes more common. The US seems to 'pick and choose' who receives its aid and at what scale due to how much the country itself will benefit. This idea roughly encompasses the concept of American exceptionalism. The term refers to the American belief that it is a morally superior country and plays a unique, unparalleled role in global politics. This belief is deeply rooted in the country's founding and associated with the spread of American ideals across the world. The idea was established with the birth of the US nation, but first appeared in full strength during World War II. America's involvement brought an end to the war and established the US as a powerful nation while increasing its nationalism. Recently, the Trump Administration has reflected back on these American exceptionalism roots and emphasizes military strength, defense, and national power. The US regards itself as a strong force that can, in theory, put a stop to most conflicts. However, this idea also supports that the US is selective in who receives its aid. Seeing as the Sudanese Civil War has not directly impacted the standard of life in the US yet, there is less of an attempt to provide as much aid as possible. Ideally, powerful countries would support the end of conflict no matter the scale, but valuable resources are hardly donated willingly, nor spread evenly. The US sees its military as one of these "valuable resources," justifying the hesitance to directly involve themselves in the Sudan conflict. Even more disheartening is the international

response to displaced civilians who are trying to seek asylum in other countries, especially the US. Specifically, in the summer of 2025, the Trump Administration deported eight men back to South Sudan after they tried to flee from the nearby conflict in Sudan. While the state department advised against returning the men back to a country in a dangerous situation, a federal judge allowed the deportation of these refugees. They have been sent back even though their country is facing the consequences of a nearby war. South Sudan is not directly involved in the fighting as it is a separate country from Sudan (since 2011), but it is exposed to many of the horrors of war. The neighboring country sees a large number of Sudanese refugees hoping to escape the violence, yet South Sudan does not have the resources to sustain many more people. People from South Sudan have been tricked and blackmailed into joining the war by the RSF, and have no way of returning back to their home country. As a regionally and historically close country to Sudan, South Sudan is exposed to the tragedies of the Sudanese Civil War. And yet, as more than one country suffers from this conflict, there is still little international aid seen.

As the US saw a shift in representation in 2024 with the presidential election, the attitude towards international aid has also changed. The Trump Administration is not interested in providing much aid to those in Sudan. Funds are being cut from USAID and may lead to the eventual shutdown of the program. President Donald Trump justified this action under the impression that the Sudanese civilians were wasting the resources and nutritional supplements sent to the country (Applebaum). People of Sudan have expressed to reporters that they haven't been wasting resources, and are even rationing it. Instead of being too much, USAID seems like too little. Some actions from the US have negative effects on the United Nations (UN). As a powerful country, the US tends to act on what's best for itself rather than the UN's opinion. As President Trump enters office once again, he proposes cutting international funding which results in an 87% decrease in UN funding (Tisdall). Even though it is not predicted that the US will leave the UN completely, it is becoming much more independent. Due to the power the US holds, other countries are raising suspicions about the UN's effectiveness. The downfall of the UN will lead to the downfall of international support, and greatly affect those in need.

These richer, more developed countries—such as the US and Russia—are able to sustain themselves without relying on international communication. Therefore, they pay less attention to countries on a different level of development or sustainability. Unfortunately these more developed and powerful areas such as North America and Western Europe have strong historical presence, especially since many territories in Africa used to be owned by European countries. During the colonization of Africa in the early 1900's, Sudan was owned and controlled by the British. The historical significance is a factor of the thriving state of these countries that Sudan and other territories in Africa will never be able to parallel due to past racism. Africa's superimposed barriers that resulted in the formation of African countries were created through racism and ideas of white supremacy. Although there has been progress and efforts to diminish racism in current times, past points of view remain influential and may be the reason for the destruction of Sudan and the UN.

The difference between aid supplied to other conflicts and the Sudanese conflict is so vast that even Sudanese civilians are noticing. A reporter from *PBS News Hour* spoke with a man who is living in a school classroom with his children. "They have forgotten us…We see people in Syria, we see people in Ukraine . . . They let us down. Now we have [gone] two months, we didn't receive anything, especially food" (PBS News Hour 22:15-22:31). The civilians express how they desperately need aid, yet are still not seeing any support. As the country spirals deeper into destruction, international aid is essential. The lack of outside help causes a loss of hope and many citizens give up on seeing an end to this war. At this point, sending weapons to either military is not as effective as sending food and health supplies to citizens. During times of hard conflict, those uninvolved yet affected require the most international attention. While a fast resolution to the overall fighting would be ideal, it is not predicted in the near future and the best course of action is harm reduction, providing care for immediate needs. Therefore, these citizens do not understand why they are being constantly overlooked.

The impact on civilians does not stop after war has ended. At this rate of destruction, researchers predict building back from war could take over a decade. The scale of the Sudanese Civil War has been observed since the start, and Sudan sees near total collapse after just one year. Now, approaching the three year mark for the conflict, the once bustling Sudan has become a graveyard. The key to helping repair Sudan lies in its citizens. More people and activists are needed to provide humanitarian aid to the country and keep its people alive. While there is no telling when and how the conflict will conclude, there are some requirements that must be met if Sudan wants to see a brighter future.

Historically, Sudan has struggled to unite its country ethnically. As a diverse community, the citizens strive for ethnic representation in the government. Due to Africa's colonization in the early 1900s, many different ethnic groups were forced together as one territory. Africa was colonized by many European powers in the 1900s, such as France and the United Kingdom. These powers did not have any consideration for the ethnic makeup of Africa, and created barriers based on the decisions agreed upon in the Berlin Conference. Since the countries in Africa were separated based on desire for land and natural resources, many ethnic groups were split up. Colonization also greatly decreased the abundance of unique ethnic, religious, and linguistic groups in Africa. Today's country barriers remain similar to those originally drawn, and Africa's populations reflect the damage of greed and colonization in the early 1900s. As a result, many countries in Africa are home to a wide variety of people who desire equal representation. They want the government to embrace the country's differences and support the community, yet this has not been seen effectively so far. If Sudan can find a way to amend its governmental policies, give more voice to its citizens, and highlight the community, it might just have a chance at rebuilding all that it has lost. Even though this political solution may help repair the government, there is little that can be done to repair the hearts and minds of those affected by the violence they have seen. Though they may see their houses fixed up and their opinions seen by the government, they will also see the makeshift gravesites scattered around villages. Citizens will not see their dead family members again after the war has concluded, and they cannot erase

the horrors they experienced. War leaves a horrific, lasting impact on people and it is likely that post-traumatic stress disorder (PTSD) will affect most, if not all, the remaining civilians of Sudan.

Conclusion

War affects people in many ways, mainly physical and psychological. The civilians who are currently experiencing the terrors of the Third Sudanese Civil War provide insight into how people are specifically impacted. They tell stories of losing hope, family, friends, and homes. On the outside, buildings in ruins and deserted towns are portrayed as the main effects on a country in war. However, the tragedies of war also impact people and the effects are much deeper and require aid during times of conflict. Physically, the death toll of the war continues to climb as families are separated and people are forced to walk across vast stretches of land in hopes of safety. These effects drive additional psychological distress. Women are raped and left with the trauma while others see their families getting murdered in front of them. Many are captured and tortured by the RSF, left to wonder if they will see their companions again. War leaves visible and invisible scars. It harms a person's identity and abandons them with long lasting trauma. The experience of living through war cannot be undone or prevented.

However, it can be helped and shortened with diplomacy and the resources of international assistance. As the United Nations (UN) is currently experiencing a reduction in power and status, however, greater awareness sparks hope of assistance and an end to the suffering of innocent people. The people of Sudan will be affected by the Sudanese Civil War for the rest of their lives while people in other countries across the world sit back and change the channel on the television, unaware of humanity's struggles.

This does not have to be the case. With aid from international powers and volunteer groups, the people of Sudan can be helped. Right now, the civilians are seeing too little. But it is not yet too late to help. The UN World Food Programme is attempting to send food and money to Sudan, but struggles with funding. The American Red Cross has established charities to raise money and help those struggling in Sudan. They focus on providing humanitarian aid such as medicine, food, and clothing instead of weapons or military resources. Islamic Relief USA (IRUSA) is another large organization that supplies help to Pakistan, Syria, Sudan, and more. IRUSA focuses on food aid, hygiene kits, medical care, shelter, and more for affected civilians. Charity Organization for Children accepts donations to help Sudan, and aims to help the young malnourished children in Sudan. They provide emergency trauma kits, mobile health clinics, and games/activities for the children as a distraction from the ongoing violence. Most of these organizations attack the hunger crisis and growing famine in Sudan due to war, and they accept online donations through their websites. These are smaller ways to provide help, yet there is the option of reaching out to a government official. While smaller branches of the government, such as a mayor or governor, may not be able to provide much help, reaching out to officials can help the issue become more recognized and even gain the attention of higher powers. When people dedicate their hearts, it is not too late to help those in need, and no attempt is too little.

Works Cited

Applebaum, Anne. "The Most Nihilistic Conflict on Earth." The Atlantic, 6 Aug. 2025, www.theatlantic.com/magazine/archive/2025/09/sudan-civil-war-humanitarian-crisis/683563/. Accessed 11 Aug. 2025.

Associated Press. "U.S. Completes Deporting 8 Men from Various Nations to South Sudan after Weeks of Legal Battles." PBS News, 5 July 2025, www.pbs.org/newshour/politics/u-s-completes-deporting-8-men-from-various-nations-to-south-sudan-after-weeks-of-legal-battles. Accessed 8 Aug. 2025.

Davies, Dave, host. "Journalist Says Greed, Nihilism and Transnationalism Are Fueling Sudan's Conflict." Africa, 6 Aug. 2025. NPR, www.npr.org/2025/08/06/nx-s1-5493499/journalist-says-greed-nihilism-and-transnationalism-are-fueling-sudans-conflict. Accessed 13 Aug. 2025.

"Inside Sudan's Civil War -- Watch the Full Documentary." YouTube, uploaded by PBS NewsHour, 16 Sept. 2024, www.youtube.com/watch?v=aSU9u-GlRMs. Accessed 22 July 2025.

"Inside Sudan's Forgotten War - BBC Africa Eye Documentary." YouTube, uploaded by BBC News Africa, 9 June 2024, www.youtube.com/watch?v=KIDMsalYHG8. Accessed 22 July 2025.

Kuol, Luka Biong D. "Civil Wars Sudan." PDF.

Mack, Burton. "A Brief History of American Exceptionalism." Yale University Press, 28 Feb. 2017, yalebooks.yale.edu/2017/02/28/a-brief-history-of-american-exceptionalism/. Accessed 14 Nov. 2025.

"On the Front Lines of Sudan's Civil War, a 'Catastrophic Situation' | WSJ." YouTube, uploaded by The Wall Street Journal, 13 Aug. 2024, youtu.be/LtLszFIFPUo?si=WAe3F-Vd2qLP4aua. Accessed 4 Aug. 2025.

Pavia, Juliet. "Power and Moral Purpose: The Rhetoric of American Exceptionalism in US Foreign Policy." Columbia University Libraries, 2023, academiccommons.columbia.edu/doi/10.7916/tnx2-nq03. Accessed 14 Nov. 2025.

Restrepo, Manuela López, host. "Who Gets Believed? Dina Nayeri Examines Society's Personal Relationship with Truth." Main Character of the Week, 10 Mar. 2023. NPR, www.npr.org/2023/03/10/1162680755/who-gets-believed-dina-nayeri-author-interview-asylum-society-truth-review. Accessed 14 Sept. 2025.

"Testimony from Frontline of Sudan's Civil War | BBC News." YouTube, uploaded by BBC News, 27 Sept. 2024, www.youtube.com/watch?v=EQFizxOZ8jQ. Accessed 4 Aug. 2025.

Tisdall, Simon. "The UN Is Our Best Defence against a Third World War. as Trump Wields the Axe, Who Will Fight to save It?" The Guardian, 6 July 2025, www.theguardian.com/commentisfree/2025/jul/06/the-un-is-our-best-defence-against-a-third-world-war-as-trump-wields-the-axe-who-will-fight-to-save-it. Accessed 22 July 2025.

"What's Happening in Sudan's Civil War? | Start Here." YouTube, uploaded by Al Jazeera English, 3 July 2024, www.youtube.com/watch?v=b12lgo_G7Hw. Accessed 22 July 2025.

Exploring the Neurobiological Basis Between Healthy and Disordered Sleep Via Sleep Patterns and Activations By Kya Li

Abstract

This literature review probes the neurobiological differences in brain activation and sleep patterns between healthy and disordered sleep. It explores the underlying pathologies of specific sleep disorders and their role in the sleep cycle and neural activation areas. Neuroimaging studies revealed that healthy sleepers exhibit regular patterns of brain activation during sleep, alternating between rapid eye movement and non-rapid eye movement sleep. This pattern helps to support optimal cognitive and emotional functioning. In contrast, sleep disorders exhibit distinct neurobiological disruptions. Disruptions in the arousal system are associated with insomnia and depression, while Parkinson's disease and narcolepsy reflect neurochemical imbalances in sleep regulation; obstructive sleep apnea and depression further demonstrate trends in overlapping neurochemical imbalances. Despite these distinctions between neurological differences, they all similarly lead to impaired brain activity and irregular sleep patterns. This review will underscore the pivotal role neurobiology plays in sleep health and pathology, drawing attention to converging patterns among sleep disorders that advance a theoretical framework for understanding shared disruptions in sleep regulation.

Introduction

Sleep is a complex neurological state whose primary function is to rest and restore energy levels in the body. Thus, sufficient sleep is crucial for maintaining health in individuals (*Eunice Kennedy Shriver National Institute of Child Health and Human Development - NICHD*). Fluctuations in the pattern of healthy sleep lead to the occurrence of sleep disorders, as seen in patients with depression, insomnia, narcolepsy, Parkinson's disease, or obstructive sleep apnea (OSA), which are detrimental to a patient's health (Kumar). Understanding the mechanisms underlying the relationship between healthy and disordered sleep is pivotal in diagnosing and preventing these disorders.

The sleep-wake cycle is an active rhythmic neural process known as the circadian rhythm. It is a brain-controlled bodily rhythm that occurs across a night's sleep and alternates between periods of sleep and wakefulness. Two major processes govern sleep-wake rhythms: the circadian process and the homeostatic process (Riemann, Nissen, et al.). Throughout one night of sleep, individuals alternate between rapid eye movement (REM) sleep – the lightest stage of sleep when dreaming occurs – and non-rapid eye movement (NREM), dreamless sleep.

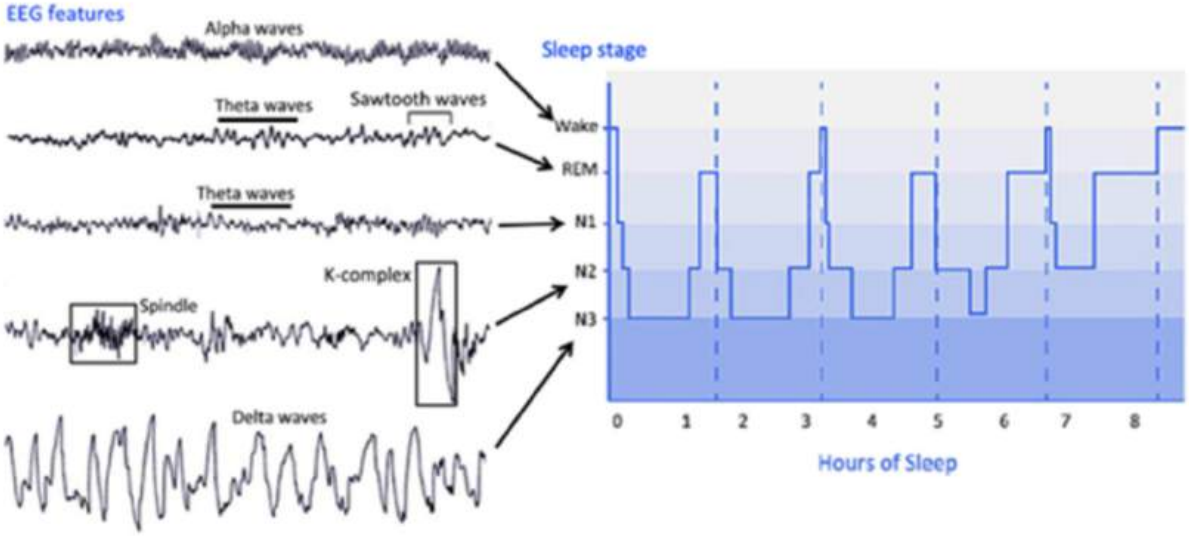


Fig 1: EEG traces and hypnogram features across sleep. Displays typical electroencephalography (EEG) features on the left, while the right shows a hypnogram visualizing sleep stage distribution over an 8-hour period. A healthy cycle of sleep depicts more deep sleep (NREM) sleep during the start of the cycle, then REM sleep prevailing in the second quarter of the night. NREM sleep includes three stages: N1 (transition to sleep), N2 (onset of sleep with spindles and K-complexes), and N3 (slow-wave sleep with delta activity and retained muscle tone) (Pan et al.).

Figure 1 provides a visual baseline for healthy sleep architecture, allowing for comparisons to be made with hypnograms from disordered sleep. For instance, insomnia is characterized by deficits in the sleep-wake cycle, where studies have found polysomnographically determined misalignments in the circadian process—visible in hypnograms as phase delays or advances—resulting in shorter sleep durations, prolonged sleep onset latency, or early morning awakenings (Palagini et al.; Van Der Woerd et al.). Depression hypnograms often resemble sleep architecture patterns similar to those in insomnia, showing fluctuating duration in sleep stages, such as reduced slow-wave sleep or fragmented sleep continuity, which contribute to non-restorative sleep and daytime fatigue (Rosenblum et al.). On the other hand, narcoleptic hypnograms span across an entire day rather than just one night's worth of sleep to demonstrate that patients constantly fall into deep stages of sleep throughout the day (Benca and Peterson; Harvard Medical School). These deviations from the healthy pattern shown in Figure 1 underscore the diagnostic value of hypnogram analysis in identifying and differentiating sleep disorders.

Like most other phenotypes, sleep can, to a certain degree, be regulated by genes. Recent research on the relationship between the brain and genetics has revealed irregularities in genes that regulate the circadian rhythm, which are prevalent in both depression and insomnia (Benca and Peterson). Additionally, studies have shown that narcolepsy may be linked to a specific

gene—particularly one associated with hypocretin (orexin) production—suggesting a genetic basis for its disruption of sleep-wake regulation.

Sleep regulation can also be influenced by neuronal systems and neurobiological patterns, with disordered sleep being understood through the role of neurotransmitters as ligands in a complex reciprocal activity (España and Scammell). These chemical disturbances are distributed across the brainstem and forebrain, each characterized by distinct projections, neurotransmitters, and receptors. For example, white matter intensities, or volumes of small lesions on brain scans that may suggest emotional and cognitive dysfunctions, were seen in patients with obstructive sleep apnea (Dauvilliers and Tafti). Whereas a loss of certain monoaminergic–transmitting messages, involving monoamine neurotransmitters and neurons in the cerebral cortex, was also seen in patients with Parkinson's disease (Wichit et al.).

Exploring the similarities and relationships between sleep disorders by analyzing evidence from past theories on neurobiological disturbances and neuroimaging techniques, this review demonstrated that sleep disorders have neurobiological bases that are more similar than previously believed (España and Scammell; Desseilles et al.; Abdrabou; Baril et al.). For instance, variation polymorphisms in genes suggest that depression and insomnia may share molecular mechanisms and biological ties. Deficiencies in specific neurotransmitters—particularly those involving hypocretin activity—have been linked to both Parkinson's disease and narcolepsy, are another example (Benca and Peterson; Fronczek et al.).

As the field of neurobiology has developed, it has become increasingly clear that there is a distinct overlap between different sleep disorders (O'Regan et al.). With dozens of sleep disorders negatively impacting sleep quality, the question arises: What future directions should be taken to elucidate such a complex phenomenon? The answer comes by developing new interpretations of the relationships between different sleep disorders. Identifying these relationships is imperative for simplifying remedies and providing fundamental insight into areas of neurobiology that are most targeted during sleep disorders by reviewing major sleep disorders with strong evidence for genetic or neuronal/chemical contributions, as they constitute the best hope for the future development of appropriate treatments in psychiatry (Benca and Peterson; Fronczek et al.; Dauvilliers and Tafti).

This literature review examines the neurobiological distinctions between healthy and disordered sleep by analyzing underlying pathologies such as brain activations and sleep patterns. Elucidating the underlying pathopsychological mechanisms of various sleep disorders offers a foundation for the development of possible therapeutic interventions.

Section I: Understanding the neurobiological bases

Insomnia

Insomnia is characterized by deficits in sleep quality and quantity. In most people, this is in regard to struggles with falling asleep or staying asleep. This sleep disorder has been hypothesized to mildly affect around 30% of the population. Within this 30%, around 10% of the

population are affected by severe or chronic insomnia (Ohayon, as cited in Fernandez-Mendoza and Vgontzas).

The hyperarousal framework has provided insights into the neurobiology of insomnia (Nofzinger et al.; Riemann, Spiegelhalder, et al.). The early work of von Economou and Moruzzi et al. led to conclusions which suggested that sleep-wake regulatory mechanisms, the ascending reticular activating system (ARAS), along with a group of hypocretin (orexin) neurons, which together maintain the switches between sleep and wakefulness to keep the sleep-wake cycle intact (Riemann, Nissen, et al.). Building on this, a model created by Saper et al. assumed that the ARAS projects towards specific brain regions: the hypothalamus, basal forebrain, and cerebral cortex, which are all areas predominantly involved in wakefulness, transmitting nervous impulses to regulate consciousness and bring sensory messages to these areas, helping individuals stay awake (Riemann, Nissen, et al.). The hypocretin neurons in the lateral hypothalamus further reinforce wakefulness by exciting various wake-promoting neurons for the physiological activation or responsiveness (Mahoney et al.). Insomnia is said to be caused when this system is disrupted, particularly through excessive ARAS activity and overabundance of hypocretin signaling. Under normal conditions, a hypothalamic "switch" can inhibit the ARAS, with hypocretin neurons helping to regulate this balance (Riemann, Spiegelhalder, et al.). A malfunctioning "switch"—due to the overactivity of the hypocretin system or hypofunction of the ventrolateral preoptic nucleus, which inhibits the ARAS— is therefore involved in the pathogenesis of primary insomnia (Riemann, Spiegelhalder, et al.). We must treat this transient sleep problem early on to prevent the spiral to chronic insomnia.

Depression

Depression manifests in a variety of symptoms that affect individuals emotionally, physically, and socially (Tracy). Around 46% of patients with major depressive disorder have experienced sleep disturbances as a common symptom - to some patients this may be displayed as an insufficiency of sleep, whereas some others may be sleeping too much (Q. Zhang et al.). Hypnograms illustrated erratic changes in the sleep cycle, particularly displaying a prolonged sleep latency and shortened total sleep time – this was caused by disturbances in slow-wave sleep (SWS). SWS is the most restorative sleep stage, where slow waves were found to be present in more than 50% of the epoch (Wichniak et al.).

Wakefulness is stimulated by cholinergic neurons, which are spread across the brain, notably in a region of the forebrain – the basal forebrain, and primarily uses the neurotransmitter acetylcholine (ACh). This neurotransmitter activates muscle contraction and enhances sensory processing and attention, firing to control REM sleep and wakefulness (André et al.). In depressed patients, elevated adenosine, a nucleoside that suppresses arousal, reduces cholinergic activity and inhibits neurons in the ventrolateral preoptic nucleus (VLPO). This then creates a causal cycle as the inhibition of ACh continues on to prevent hypothalamic hypocretin-containing neurons, as well as activating systems mentioned previously from firing (Liu and Gao; Ziganshin et al.).

Neurons of the lateral hypothalamus release hypocretin during wakefulness, stimulating the action of monoaminergic systems: for example, serotonergic neurons found in the dorsal raphe nuclei (dRN), and the noradrenergic neurons in the locus coeruleus (LC) (Wichniak et al.). This activity supports transitions from wakefulness to sleep (Dalal et al.). The reciprocal-interaction model suggests that brainstem neurons, particularly those in the pons and the medulla, regulate sleep stages by exchanging electrical signals and chemicals through the brain (Dalal et al.). These interactions are either REM-on, which fires during REM sleep, or REM-off, firing during wakefulness but not REM sleep (Schwartz and Kilduff). This model gained attention for its potential to explain the cause of depression in sleep, highlighting REM-on cholinergic neurons in the midbrain region of the brainstem, but monoaminergic REM-off cells in the LC and the dRN, which stimulated wakefulness (McCarley).

This view was further supported by pharmacological evidence: antidepressants enhance monoaminergic neurotransmission, providing therapeutic benefit for depressed patients (Morilak and Frazer). Monoamine neurons play a crucial role in regulating mood – an example of this is serotonin (5-HT), which helps regulate positive mood and emotion. Accordingly, we could interpret that a lack of monoaminergic neurons could lead to depression (Thase; Wichniak et al.).

Parkinson's disease

Parkinson's disease (PD) is a chronic and progressive disorder that primarily affects motor control due to the loss of dopaminergic neurons in the brain (Zhou et al.). In addition to core motor symptoms such as rigidity, postural instability, and tremors, PD can also significantly disrupt sleep (Thangaleela et al.). Electrosomnograms show that patients with PD experience disturbed sleep spindles and K-complexes, or insufficient muscle atonia (lack of normal muscular tension). As a result, common symptoms include nocturnal sleep disturbances and excessive daytime sleepiness (EDS) (Lima). A standardized paradigm known as the Multiple Sleep Latency Test (MSLT) found that around 40% of PD patients have mean sleep latencies of approximately 5 minutes, thus providing unequivocal evidence of EDS (Bargiotas et al.).

As dopamine is widely associated with wakefulness, an imbalance in dopamine levels may also contribute to EDS. The primary neurobiological issue in PD is the degeneration of dopaminergic neurons. Dopamine modulates behaviors such as facilitating smooth and coordinated movement and the maintenance of wakefulness (Eban-Rothschild et al.). Fluctuations in dopaminergic neurotransmission can therefore lead to exhaustion, weariness, and irregular movements, common characteristics of PD. As these neurons degenerate, dopamine levels in the brain decrease significantly, leading to a decrease in dopamine signals from the basal ganglia to the motor cortex (Ramesh and Arachchige). This demonstrates that dopamine deficits result in dysfunctions of neuronal circuits responsible for controlling motor execution, particularly within the basal ganglia – a brain region innervated by dopaminergic afferents involved in motor function (Zhou et al.). Animal models of PD using dopaminergic neurotoxins (chemical compounds that selectively damage dopamine-producing neurons) have demonstrated

sleep-related non-motor symptoms even without dopaminergic treatment, emphasizing the intrinsic link between dopamine loss and sleep-wake dysregulation (Dovonou et al.).

Narcolepsy

Narcolepsy is a sleep disorder that exhibits an intrusion of REM sleep during wakefulness, EDS, episodes of muscle paralysis and atonia induced by cataplexy, as well as hypnagogic (immediately after waking) and hypnopompic (when falling asleep) hallucinations (Mahoney et al.).

Hypocretin neurons are localized in the hypothalamus and have widespread projections throughout the brain. Hypocretins are neuropeptides that play a role in maintaining wakefulness and orchestrating REM sleep physiology. Experiments with mice have supported this idea, finding that hypocretin helps to generate a high arousal state and allows an animal to become attentive and primed for action. Loss of hypocretin therefore reduces attentiveness and arousal, leading to narcolepsy (Hauw et al.). Particularly in narcolepsy type 1 (NT1), the loss of hypocretin neurons in the hypothalamus disrupts arousal pathways, contributing to impaired wakefulness (Mahoney et al.).

Dysfunctions in neuroanatomical structures also contribute to narcolepsy. It has been hypothesized that the neuropathological basis of narcolepsy may be explained by a significant decrease in gray matter concentration in specific brain regions (Desseilles et al.; Weng et al.). This reduction compromises neuronal health, leading to diminished synaptic activity, and thus eliminates the excitatory drive from hypocretin neurons. Consequently, signals from the amygdala inhibit REM sleep-suppressing regions, enabling activity in the sublaterodorsal nucleus (SLD)—which drives muscle paralysis—and thus resulting in symptoms of narcolepsy (Mahoney et al.).

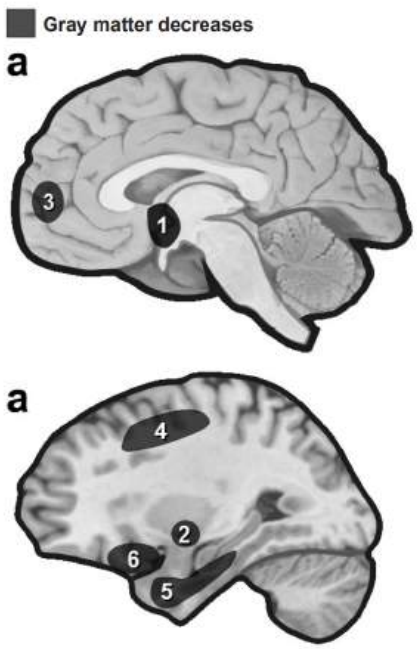


Fig 2: Gray matter deficits in narcoleptic patients: a voxel-based morphometry (VBM) study. Anatomical brain changes in narcoleptic patients assessed by VBM in the frontal brain regions of narcoleptic patients. Hyperperfusion, an abnormal increase in blood flow, has had a close association with gray matter atrophy across several brain regions. Namely, the frontomesial (3), the hypothalamus (1), the right prefrontal cortex (4), right nucleus accumbens (2), inferior frontal cortex (6), and the inferior temporal cortex (5) (Desseilles et al.).

After observing 20 narcoleptic patients and comparing their neuroanatomy with healthy controls, Figure 2 highlights that these regions were most affected by gray matter loss (Desseilles et al.). Later research discovered diminished volume in the thalamus and amygdala—limbic structures not systematically detected by VBM due to their small size and susceptibility to imaging artifacts (Weng et al.).

Importantly, this decrease in gray matter may represent the secondary neuronal losses of hypocretin neurons, widely considered the root cause of narcolepsy. Such structural deficits provide evidence that narcolepsy is characterized by both functional dysregulation and anatomical degeneration, linking hypocretin deficiency to broader impairments in sleep–wake regulation and emotional processing.

Obstructive sleep apnea

Sleep apnea has been characterized by reductions or cessations in airflow associated with partial or complete obstruction of the upper airway during sleep. Obstructive sleep apnea (OSA) is a common type of sleep apnea. Key symptoms of OSA include abrupt awakenings accompanied by gasping or choking, EDS, early morning headaches, and irregular breathing patterns (Patil et al.).

In OSA patients, structural and functional changes on brain imaging using functional magnetic resonance imaging (fMRI) have been previously associated with neurocognitive

deficits and reduced functional connectivity in specific projecting fibers of the brain (Wu et al.). Gray matter reductions have been observed in the right temporal gyrus and cerebellum (Baril et al.), as well as in the amygdala, insular cortex, and the hippocampus (Philby et al.). Because the cerebellum controls coordination and movement, atrophy of gray matter in this region may result in loss of muscle coordination. A meta-analysis of 15 publications including 290 OSA patients and 290 controls reported that there was a convergence for structural and functional differences in OSA patients of gray matter reduction and functional disturbance in the right basolateral nucleus of the amygdala/hippocampus and the right central insula (Rosenzweig et al.).

The severity of the sleep apnea may also lead to intermittent hypoxia, a condition whereby an individual has abnormally low levels of oxygen in their bodily tissues due to the blockage of the upper airway. Hypoxia induces chronic inflammation, which can damage neurons and connections, contributing to neurodegeneration, disturbed neural integrity, and gray matter atrophy (Mukandala et al.). It also impairs synaptic plasticity – the ability of synapses to strengthen or weaken over time, and to an extent, affect neurogenesis – the formation of new neurons, which both aim to maintain gray matter volume (McKenna et al.).

White matter hyperintensities (WMH) can also be a leading cause of OSA. It represents the deterioration and damage to white matter in the brain, and drastically reduces cognitive abilities (Schammel et al.). A non-invasive imaging technique called Diffusion Tensor Imaging (DTI), which measures neuronal structures and highlights the subtle changes to white matter, helped explain the significance of WMH to OSA (Abdrabou). Biomarkers of white matter microstructure in diffusion MRI studies are analyzed in four different ways:

- Fractional Anisotropy (FA), measuring the degree of directionality of water diffusion
- Mean Diffusivity (MD) measures the average rate of water diffusion in all directions within the brain's white matter
- Radial Diffusivity (RD), measuring the diffusion of water perpendicular to the main direction of the white matter fibers
- Axial Diffusivity (AD), which measures the magnitude of the diffusion following white matter fibers (Baril et al.).

OSA patients demonstrated lower FA and higher MD in the anterior corpus callosum, a white matter tract connecting the frontal lobes. Increased RD suggested myelin sheath damage, impairing nerve conduction and leading to demyelination. These abnormalities promote WMH accumulation (Baril et al.), impacting interhemispheric communication and potentially affecting motor functions and increasing the likelihood of conditions such as cognitive decline or strokes. A study by Carvalho et al. conducted on 140 individuals with OSA also concluded that severe sleep apnea patients displayed a higher volume of white matter hyperintensities and reduced axonal integrity (Dauvilliers and Tafti).

These five sleep disorders are associated with impaired health and disrupted sleep quality and duration, each manifested through distinct neurobiological changes. Some involve alterations in sleep patterns, whereas others focus on damage to brain activation networks (Gottesman et al.). The neurobiological underpinnings of insomnia are illuminated by the hyperarousal process framework. This framework posits that dysfunctions in the ARAS and an abnormal abundance of hypocretin neurons, which normally maintain consciousness and promote wakefulness, will lead to difficulties initiating or maintaining sleep, resulting in insomnia (Jones; Van Someren). Depression also reflects on erratic neural activation patterns within the neurotransmitter systems; cholinergic neurons, which control alertness and REM sleep, are suppressed in depressed patients due to the irregular overactivity of adenosine. Another factor that may arouse depression is having REM-off in the locus coeruleus and the dorsal raphe nuclei. These areas trigger the serotonergic and the noradrenergic systems. If these systems are REM-off, it explains why depressed patients may struggle with falling asleep (Xie et al.). PD introduces further complexity, wherein the degeneration of dopaminergic neurons contributes not only to motor dysfunction but also to disturbances in sleep architecture, characterized by excessive daytime sleepiness (EDS). Similarities were seen in narcolepsy, particularly type 1, from results regarding hypocretin neuron loss, which is critical for maintaining arousal and regulating REM sleep, as well as gray matter loss in areas of the brain. While obstructive sleep apnea (OSA) leads to significant neurocognitive deficits and alterations in brain structure associated with gray matter loss and WMH (Joo et al.). Furthermore, potential damage to the myelin sheath (a feature of the axon) hinders the transmission of nerve signals and may mean that chemical messengers are unable to reach the upper airway duct, leading to its blockage (Kumar et al.). Collectively, these conditions underscore the intricate interrelation between sleep regulation mechanisms and neurobiological health, indicative of the profound impact sleep disorders impose on emotional and cognitive functioning.

Section II: Exploring relationships

Insomnia and Depression

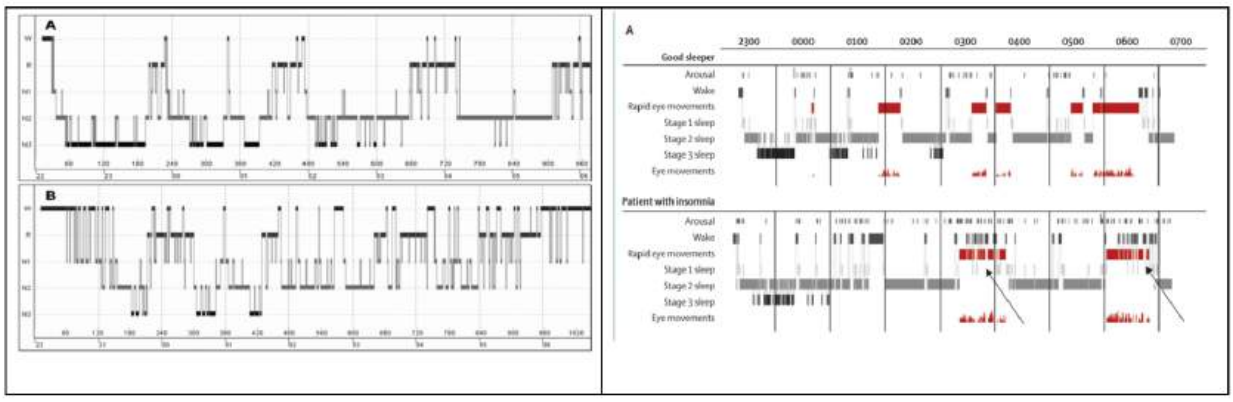


Fig 3: Hypnograms displaying sleep patterns in patients with depression versus insomnia. Side-by-side hypnograms compare normal sleep patterns architecture against the altered sleep patterns seen in these disorders. The primary

disturbances arise from frequent stage shifts and increased brief waking periods. At a chronobiological level, these disruptions may be associated with misalignments in the circadian process, leading to phase advances or delays (Wichniak et al.; Riemann, Nissen, et al.).

As shown in Figure 3, the hypnograms of patients with depression and insomnia reveal strikingly similar patterns of sleep fragmentation and instability. These parallels found between depression and insomnia hypnograms thus play a role in delineating which sleep disorders may be closely linked together and share similar neurobiological bases. Insomnia has been identified as the magnification or the development of other psychiatric disorders, such as depression (Riemann, Nissen, et al.). Around 14% of patients with persistent insomnia had concurrent depression, whereas depression occurred in less than 1% of patients with no sleep complaints (Ford and Kamerow, as cited in Roth). The strong correlation between the two sleep disorders may therefore provide an explanation as to why diagnostic evaluations of insomnia should account for depressive symptoms and neurobiology, and likewise, assessments of depression should consider characteristics of insomnia.

a) Brain activations

The ARAS is fundamental to wakefulness, and dysfunction within the hypothalamus can slow down ARAS activity, the mechanism that underlies insomnia. The ARAS is influenced by various chemical impulses, each mediated by specific neurotransmitters and pathways, with serotonin being one such neurotransmitter (Jones). A functional decrease of serotonergic neurotransmission contributes to depression and is associated with alterations of sleep, notably insomnia. Lesion studies have shown that serotonergic raphe neurons, types of serotonergic neurons specified in a region called the raphe nuclei, are not only responsive in regulating mood but also behavioral arousal (McGinty and Harper; Kiyasova and Gaspar). Since the ARAS sends projections to the thalamus, hypothalamus, and the cortex, it regulates the firing of neurons in the raphe nuclei. Malfunction of the ARAS can therefore indirectly alter serotonergic activity through its connections with these regions. Such dysfunction hinders the regulation of serotonergic raphe neurons, disrupting both mood regulation and arousal, leading to the likelihood of both depression and insomnia (Jones).

b) Activation of the HPA-axis

The overactivity of the hypothalamic-pituitary-adrenal (HPA) axis has been observed in both insomnia and depression. The HPA axis stimulates cortisol secretion – a hormone which mediates responses to the ‘fight or flight’ response (Nicolaides et al.). Abnormally elevated cortisol levels interfere with the body's sleep-wake cycle and lead to difficulties with sleep initiation, maintenance of sleep, or achieving restorative sleep. Under normal conditions of sleep, cortical secretion enhances REM sleep, whereas SWS has an inhibitory influence on the HPA axis and cortical secretion, leading to arousal and sleeplessness (Keller et al.; Elder et al.).

A study conducted by Vgontzas et al. (1998), later supported by polysomnographic recordings from Rodenbeck et al. (2002), suggested that the overactivity of the HPA-axis and

high cortisol levels were present in insomniacs with short sleep duration. Similar studies found that cortisol values were elevated by 15–20% in the insomnia group compared with a control group (Vgontzas et al.; Rodenbeck et al.). However, opposing evidence was found from Riemann et al. (2002), suggesting there were no cortisol differences observed between the sleep of patients with insomnia and controls (Riemann, Klein, et al.). This suggests that an extensive amount of research is needed to evaluate cortisol excretion in regard to sleep disorders, as results collected remain inconclusive.

HPA dysfunction has also been implicated in psychotic depression. Non-suppression occurs when cortisol levels remain elevated despite the administration of dexamethasone, a steroid that supposedly helps to suppress cortisol and attempts to regulate the body's circadian periodicity. Patients with psychotic depression often show high rates of non-suppression in cortisol on the dexamethasone suppression test (DST), and elevated cortisol levels persisting post-dexamethasone, confirming that the irregularities and dysfunction in the HPA axis can be seen in depression. Furthermore, neuroimaging revealed that patients with recurrent depression exhibited evidence of hippocampal atrophy (degeneration of cells), which correlates strongly with the total duration of depressive episodes – supporting the idea that prolonged exposure to elevated cortisol contributes to structural brain changes. This idea thus appears independent of aging and reflects neurotoxic changes linked to HPA-axis dysregulation (Sheline et al.).

c) Genetics

Sleep and sleep disorders are complex phenotypes that are regulated by many genes, thus making it extremely difficult to find a clear-cut single gene relationship, but delineating one or several genes that may be predictive of a certain disorder will offer significant steps forward in indications. For example, irregularities in the circadian locomotor output cycles of the kaput (CLOCK) gene have been demonstrated to affect both depression and insomnia (Benca and Peterson), with another study demonstrating that patients with a depressive disorder who have a variant polymorphism in their CLOCK gene were more likely to experience lifetime insomnia (Serretti et al.).

From the research collected above, it is likely that an individual will be prone to both disorders because of some similarities in pathology. We can conclude that insomnia and depression may account for similar treatments, and monitoring depression symptoms in insomnia patients may help identify depression early in its evolution.

Narcolepsy and PD

a) Hypocretin neurotransmission and CSF

The relationship between narcolepsy and PD was first proposed due to their similarities in symptoms, particularly EDS, frequent sleep attacks, and muscle atonia. When normal sleep-wake processes are disturbed, for instance, through an extensive loss of neuronal projections, neurotransmitters would be inhibited (Jones). Research has reported that individuals

with narcolepsy or PD may both experience a diminution of hypocretin neurons during hypocretin neurotransmission.

In NT1, hypocretin-producing neurons in the hypothalamus are lost, resulting in low levels of hypocretin in the cerebrospinal fluid (CSF). This is significant, as levels of CSF hypocretin have been theorized by the International Classification of Sleep Disorders (ICSD-2) to be a diagnostic test for narcolepsy. CSF facilitates neural communication by increasing efficiency and regulating brain activity patterns and sleep architecture, as well as promoting restorative processes and indirectly helping to support sleep-wake regulation (Fronczek et al.), which suggests why low levels of this fluid may lead to symptoms of narcolepsy. Supporting this, Nishino et al. concluded that low levels of CSF hypocretin exist in human narcoleptics, having an average of around ≤110 pg/mL compared to healthy controls, whose levels averaged around 300.5 pg/mL.

NT1 tissue concentrations were also found to be reduced in patients with PD. Hypocretin concentrations in the prefrontal cortex were almost 40% lower in PD patients compared with controls, while CSF levels showed a reduction of about 25%. A French study conducted in 2003 reported significantly reduced levels of CSF in PD patients (57 pg/ml) compared with controls (199 pg/ml) (Haq et al.). In addition, post-mortem analyses of ventricular CSF levels by Fronczek et al. revealed a 20% reduction of hypocretin-1 in PD patients relative to controls.

However, Thannickal et al suggested that the destruction of cells in the hypothalamus by PD was not selective to just hypocretin neurons, as they showed a degeneration of other cell types (melanin-concentrating hormone) alongside hypocretin cells. Therefore, although the function of CSF has been previously associated with narcolepsy, lower levels of CSF hypocretin in PD patients are still debated (Thannickal et al.).

Obstructive sleep apnea and depression

The association between obstructive sleep apnea (OSA) and depression has been widely investigated, with studies reporting both supporting and conflicting evidence. 45% of 55 OSA patients had depressive symptoms on the Zung Self-Rating Depression Scale. 26% of OSA patients described themselves as depressed, while 58% of individuals fulfilled the Diagnostic and Statistical Manual of Mental Disorders, third edition (DSM-III) criteria for severe depression, having four or more depression symptoms (Schröder and O'Hara). However, there is also evidence that refutes the hypothesis that there is a relationship between OSA and depression. In Schroder and O'Hara's study, they cite Philips et al.'s longitudinal study stating that depressive symptoms in elderly patients with mild-OSA were unestablished, and relationships between the two were obscure.

a) Sleep fragmentation

One mechanism that may explain the association between OSA and depression is the role of EDS. EDS is a recurrent symptom of OSA but is rarely observed in patients with depression alone (D. Zhang et al.). But recent studies have suggested that EDS in OSA patients can result in

depression symptomatology in OSA, as persistent sleepiness exacerbates mood disturbances and intensifies depressive features (Jehan et al.). Reliable measures such as the Epworth Sleepiness Scale (ESS), a self-administered questionnaire that asks respondents to rate their likelihood of falling asleep in everyday life encounters, and the Maintenance of Wakefulness Test (MWT), a clinical assessment used to evaluate daytime alertness and wakefulness, have demonstrated that EDS correlates with higher depression scores on the Hospital Anxiety and Depression Scale (HAD-D) (Cai et al.; Lachowicz et al.).

Neuroimaging evidence further supports this link. Individuals with OSA consistently exhibit structural brain changes, namely reductions in gray matter volume and increased WMH, particularly in regions responsible for cognition and mood regulation. These structural changes from intermittent hypoxia and sleep fragmentation may impair cognition and mood regulation, which are not only symptoms commonly observed in individuals with OSA, but such characteristics can also increase vulnerability to depression. Thus it is possible that such impairments in brain structures could lead to depression as well. Aloia et al. found a tendency for gray and white subcortical hyperintensities to be positively correlated with depression scores on the Hamilton Depression Scale. Therefore, gray matter damage or WMH will likely not only lead to OSA but also mood disorders, which may lead to depression (Aloia et al.; Zacharias et al.; Baril et al.).

b) Upper airway control and the role of serotonin

Another shared mechanism between OSA and depression involves serotonergic signaling. The serotonergic system has a central role in regulating mood and the sleep-wake cycle, but it also influences upper airway muscle tone during sleep. By binding to serotonin receptors, it will influence (enhance or inhibit) the activity of motor neurons. Enhanced activity of the upper airway dilator muscles to keep the airway open for respiration, whereas inhibitory effects reduce muscle tone and increase the risk of airway collapse. With a lack of serotonergic neurons, delivery to the upper airway dilatator motor neurons will be reduced, contributing to OSA. At the same time, serotonergic dysfunction impairs mood regulation, highlighting a shared neurochemical mechanism between the two conditions (Witkowska et al.).

Pharmacological evidence reinforces this connection. Antidepressants such as selective serotonin reuptake inhibitors (SSRIs), which aim to increase serotonergic neurotransmission, are widely prescribed molecules that aim to relieve feelings of depression (Schröder and O'Hara). The antidepressants were suggested to similarly improve the apnea-hypopnea index (AHI), a measure used to assess the severity of OSA. Moreover, several serotonin receptor ligands and bi-functional molecules are under development, with the potential to simultaneously target depressive syndromes and OSA (Caples et al.).

Discussion

Research into sleep disorders remains constrained by the heterogeneity of methodologies employed in neuroimaging studies. For example, small sample sizes, variability in disease

duration, demographic differences, gender biases, and insufficient consideration of biopsychosocial factors all reduce the generalizability of findings. Publication and author biases also further complicate interpretation. Many studies also fail to account for different stages or severities of the same disorder. For example, the causations of acute insomnia may differ from those of chronic insomnia; it remains unclear whether hyperarousal is unique to primary insomnia or if it is comorbid with other psychiatric disorders.

Research on narcolepsy within the literature review faces challenges as results have often produced equivocal results, failing to distinguish between narcolepsy with cataplexy (NT1) or narcolepsy without cataplexy (NT2). While NT1 has been more clearly linked to hypocretin deficiency, the pathophysiology of NT2 remains elusive, as hypocretin levels remain normal. This suggests that NT2 may arise from entirely different mechanisms, making it unsuitable for clarifying narcolepsy neuropathology. Neuroimaging may therefore require improved resolution, sensitivity, and specificity in imaging hardware to uncover subtle brain changes that current techniques often miss, such as gray matter destruction in narcolepsy, gray matter atrophy in narcolepsy, WMH in OSA, or neurotransmitter loss in depression and PD.

Addressing these limitations will require larger, diverse populations and standardized protocols to better capture activation patterns and underlying mechanisms across sleep disorders. At the same time, therapeutic approaches must evolve to reflect the complex and bidirectional relationships between neurobiology and clinical symptoms. For example, dysfunction of serotonin neurotransmission may not be the primary cause of depression; rather, depression itself may lead to neurotransmitter deficits. Recognizing such reciprocal dynamics is essential for developing interventions that target root causes of sleep disorders rather than merely alleviating symptoms. As neurobiology advances, new discoveries can quickly reshape accepted ideas, underscoring the need for adaptable approaches that integrate emerging evidence into clinical practice.

Conclusion

Reviewing the extensive literature on the neurobiological basis of sleep disorders reveals the key findings that illuminate the complex interplay between brain activations and sleep patterns, concluding that sleep disorders may cause anomalies in the sleep-wake cycle. This disruption affects brain activation and neural patterns and stems from genetic predispositions and imbalances in neurotransmitter systems. Using VBM, fMRI, and MRI neuroimaging studies, along with meta-analyses and whole-brain studies, we concluded definitive relationships between sleep disorders. This offers a provisional step towards a theoretical framework to guide initial steps that could be taken to inform possible treatments for more than one sleep disorder.

Works Cited

Abdrabou, Ahmed. "Diffusion Tensor Imaging and Fiber Tractography | Radiology Reference Article | Radiopaedia.Org." Radiopaedia, https://doi.org/10.53347/rID-22702. Accessed 30 Aug. 2025.

Aloia, Mark S., et al. "Neuropsychological Sequelae of Obstructive Sleep Apnea-Hypopnea Syndrome: A Critical Review." Journal of the International Neuropsychological Society, vol. 10, no. 5, Sept. 2004, pp. 772–85. Cambridge University Press, https://doi.org/10.1017/S1355617704105134.

André, Claire, et al. "REM Sleep Is Associated with the Volume of the Cholinergic Basal Forebrain in aMCI Individuals." Alzheimer's Research & Therapy, vol. 15, no. 1, Sept. 2023, p. 151. BioMed Central, https://doi.org/10.1186/s13195-023-01265-y.

Bargiotas, Panagiotis, et al. "Sleepiness and Sleepiness Perception in Patients with Parkinson's Disease: A Clinical and Electrophysiological Study." Sleep, vol. 42, no. 4, Apr. 2019, p. zsz004. PubMed, https://doi.org/10.1093/sleep/zsz004.

Baril, Andrée-Ann, et al. "Obstructive Sleep Apnea and the Brain: A Focus on Gray and White Matter Structure." Current Neurology and Neuroscience Reports, vol. 21, no. 3, Feb. 2021, p. 11. Springer Link, https://doi.org/10.1007/s11910-021-01094-2.

Benca, Ruth M., and Michael J. Peterson. "Insomnia and Depression." Sleep Medicine, vol. 9, Sept. 2008, pp. S3–9. DOI.org (Crossref), https://doi.org/10.1016/S1389-9457(08)70010-8.

Cai, Si-Jie, et al. "Correlation of Epworth Sleepiness Scale with Multiple Sleep Latency Test and Its Diagnostic Accuracy in Assessing Excessive Daytime Sleepiness in Patients with Obstructive Sleep Apnea Hypopnea Syndrome." Chinese Medical Journal, vol. 126, no. 17, 2013, pp. 3245–50.

Caples, Sean M., et al. "Obstructive Sleep Apnea." Annals of Internal Medicine, vol. 142, no. 3, Feb. 2005, pp. 187–97. PubMed, https://doi.org/10.7326/0003-4819-142-3-200502010-00010.

Dalal, Jasbir, et al. "Translational Profiling of Hypocretin Neurons Identifies Candidate Molecules for Sleep Regulation." Genes & Development, vol. 27, no. 5, Mar. 2013, pp. 565–78. PubMed, https://doi.org/10.1101/gad.207654.112.

Dauvilliers, Yves, and Mehdi Tafti. "The Genetic Basis of Sleep Disorders." Current Pharmaceutical Design, vol. 14, no. 32, 2008, pp. 3386–95. PubMed, https://doi.org/10.2174/138161208786549308.

Desseilles, Martin, et al. "Neuroimaging Insights into the Pathophysiology of Sleep Disorders." Sleep: Journal of Sleep and Sleep Disorders Research [US], vol. 31, no. 6, 2008, pp. 777–94. APA PsycNet, https://doi.org/10.1093/sleep/31.6.777.

Dovonou, Axelle, et al. "Animal Models of Parkinson's Disease: Bridging the Gap between Disease Hallmarks and Research Questions." Translational Neurodegeneration, vol. 12, no. 1, July 2023, p. 36. BioMed Central, https://doi.org/10.1186/s40035-023-00368-8.

Eban-Rothschild, Ada, et al. "VTA Dopaminergic Neurons Regulate Ethologically Relevant Sleep–Wake Behaviors." Nature Neuroscience, vol. 19, no. 10, Oct. 2016, pp. 1356–66. PubMed Central, https://doi.org/10.1038/nn.4377.

Elder, Greg J., et al. "Stress and the Hypothalamic–Pituitary–Adrenal Axis: How Can the COVID-19 Pandemic Inform Our Understanding and Treatment of Acute Insomnia?"

Journal of Sleep Research, vol. 32, no. 4, 2023, p. e13842. Wiley Online Library, https://doi.org/10.1111/jsr.13842.

España, Rodrigo A., and Thomas E. Scammell. “Sleep Neurobiology for the Clinician.” Sleep, vol. 27, no. 4, June 2004, pp. 811–20.

Eunice Kennedy Shriver National Institute of Child Health and Human Development - NICHD. https://www.nichd.nih.gov/health/topics/factsheets/sleep. Accessed 30 Aug. 2025.

Fernandez-Mendoza, Julio, and Alexandros N. Vgontzas. “Insomnia and Its Impact on Physical and Mental Health.” Current Psychiatry Reports, vol. 15, no. 12, Dec. 2013, p. 418. PubMed Central, https://doi.org/10.1007/s11920-013-0418-8.

Fronczek, Rolf, et al. “Hypocretin (Orexin) Loss in Parkinson’s Disease.” Brain, vol. 130, no. 6, June 2007, pp. 1577–85. Silverchair, https://doi.org/10.1093/brain/awm090.

Gottesman, Rebecca F., et al. “Impact of Sleep Disorders and Disturbed Sleep on Brain Health: A Scientific Statement From the American Heart Association.” Stroke, vol. 55, no. 3, Mar. 2024, pp. e61–76. ahajournals.org (Atypon), https://doi.org/10.1161/STR.0000000000000453.

Haq, Imran Z, et al. “Narcolepsy in Parkinson’s Disease.” Expert Review of Neurotherapeutics, vol. 10, no. 6, June 2010, pp. 879–84. Taylor and Francis+NEJM, https://doi.org/10.1586/ern.10.56.

Harvard Medical School. Understanding Narcolepsy | Sleep Medicine. https://sleep.hms.harvard.edu/education-training/public-education/sleep-and-health-education-program/sleep-health-education-2. Accessed 30 Aug. 2025.

Hauw, Jean-Jacques, et al. “Neuropathology of Sleep Disorders: A Review.” Journal of Neuropathology & Experimental Neurology, vol. 70, no. 4, Apr. 2011, pp. 243–52. Silverchair, https://doi.org/10.1097/NEN.0b013e318211488e.

Jehan, Shazia, et al. “Depression, Obstructive Sleep Apnea and Psychosocial Health.” Sleep Medicine and Disorders : International Journal, vol. 1, no. 3, 2017, p. 00012.

Jones, Barbara E. “Chapter 9 - Neurobiology of Waking and Sleeping.” Handbook of Clinical Neurology, edited by Pasquale Montagna and Sudhansu Chokroverty, vol. 98, Elsevier, 2011, pp. 131–49. Sleep Disorders Part I. ScienceDirect, https://doi.org/10.1016/B978-0-444-52006-7.00009-5.

Joo, Eun Yeon, et al. “Reduced Brain Gray Matter Concentration in Patients With Obstructive Sleep Apnea Syndrome.” Sleep, vol. 33, no. 2, Feb. 2010, pp. 235–41. PubMed Central, https://doi.org/10.1093/sleep/33.2.235.

Keller, J., et al. “HPA Axis in Major Depression: Cortisol, Clinical Symptomatology and Genetic Variation Predict Cognition.” Molecular Psychiatry, vol. 22, no. 4, Apr. 2017, pp. 527–36. www.nature.com, https://doi.org/10.1038/mp.2016.120.

Kiyasova, Vera, and Patricia Gaspar. “Development of Raphe Serotonin Neurons from Specification to Guidance.” The European Journal of Neuroscience, vol. 34, no. 10, Nov. 2011, pp. 1553–62. PubMed, https://doi.org/10.1111/j.1460-9568.2011.07910.x.

Kumar, Rajesh, et al. “Abnormal Myelin and Axonal Integrity in Recently Diagnosed Patients with Obstructive Sleep Apnea.” Sleep, vol. 37, no. 4, Apr. 2014, pp. 723–32. PubMed Central, https://doi.org/10.5665/sleep.3578.

Kumar, Velayudhan Mohan. “Sleep and Sleep Disorders.” The Indian Journal of Chest Diseases & Allied Sciences, vol. 50, no. 1, 2008, pp. 129–35.

Lachowicz, J, et al. “O045 Anxiety and Daytime Sleepiness Scores Have a Weak but Statistically Significant Correlation with Apnoea-Hypopnoea Index.” SLEEP Advances, vol. 2, no. Supplement_1, Oct. 2021, p. A20. Silverchair, https://doi.org/10.1093/sleepadvances/zpab014.044.

Lima, Marcelo M. S. “Sleep Disturbances in Parkinson’s Disease: The Contribution of Dopamine in REM Sleep Regulation.” Sleep Medicine Reviews, vol. 17, no. 5, Oct. 2013, pp. 367–75. ScienceDirect, https://doi.org/10.1016/j.smrv.2012.10.006.

Liu, Zhong-Wu, and Xiao-Bing Gao. “Adenosine Inhibits Activity of Hypocretin/Orexin Neurons by the A1 Receptor in the Lateral Hypothalamus: A Possible Sleep-Promoting Effect.” Scilit, 1 Jan. 2007, https://www.scilit.com/publications/a78570a07deef389fc6cfcdf9927dceb.

Mahoney, Carrie E., et al. “The Neurobiological Basis of Narcolepsy.” Nature Reviews Neuroscience, vol. 20, no. 2, Feb. 2019, pp. 83–93. www.nature.com, https://doi.org/10.1038/s41583-018-0097-x.

McCarley, Robert W. “REM Sleep and Depression: Common Neurobiological Control Mechanisms.” The American Journal of Psychiatry [US], vol. 139, no. 5, 1982, pp. 565–70. APA PsycNet, https://doi.org/10.1176/ajp.139.5.565.

McGinty, Dennis J., and Ronald M. Harper. “Dorsal Raphe Neurons: Depression of Firing during Sleep in Cats.” Brain Research [Netherlands], vol. 101, no. 3, 1976, pp. 569–75. APA PsycNet, https://doi.org/10.1016/0006-8993(76)90480-7.

McKenna, James T., et al. “Aberrant Cortical Neuroplasticity in the OSA Patient (Commentary on Opie et Al.).” European Journal of Neuroscience, vol. 37, no. 11, 2013, pp. 1843–1843. Wiley Online Library, https://doi.org/10.1111/ejn.12221.

Morilak, David A., and Alan Frazer. “Antidepressants and Brain Monoaminergic Systems: A Dimensional Approach to Understanding Their Behavioural Effects in Depression and Anxiety Disorders.” The International Journal of Neuropsychopharmacology, vol. 7, no. 2, June 2004, pp. 193–218. PubMed, https://doi.org/10.1017/S1461145704004080.

Mukandala, Gatambwa, et al. “The Effects of Hypoxia and Inflammation on Synaptic Signaling in the CNS.” Brain Sciences, vol. 6, no. 1, Feb. 2016, p. 6. PubMed Central, https://doi.org/10.3390/brainsci6010006.

Nicolaides, Nicolas C., et al. “HPA Axis and Sleep.” Endotext, edited by Kenneth R. Feingold et al., MDText.com, Inc., 2000. PubMed, http://www.ncbi.nlm.nih.gov/books/NBK279071/.

Nofzinger, Eric A., et al. “Functional Neuroimaging Evidence for Hyperarousal in Insomnia.” American Journal of Psychiatry, vol. 161, no. 11, Nov. 2004, pp. 2126–28. psychiatryonline.org (Atypon), https://doi.org/10.1176/appi.ajp.161.11.2126.

Ohayon, Maurice M. “Epidemiology of Insomnia: What We Know and What We Still Need to Learn.” Sleep Medicine Reviews, vol. 6, no. 2, Apr. 2002, pp. 97–111. PubMed, https://doi.org/10.1053/smrv.2002.0186.

O’Regan, David, et al. “Sleep Neurophysiology and Psychiatric Disorders: A Transdiagnostic Framework for Mechanistic and Therapeutic Insight.” Current Treatment Options in Neurology, vol. 27, no. 1, Aug. 2025, p. 36. Springer Link, https://doi.org/10.1007/s11940-025-00847-5.

Palagini, Laura, et al. “Insomnia, Sleep Loss, and Circadian Sleep Disturbances in Mood Disorders: A Pathway toward Neurodegeneration and Neuroprogression? A Theoretical Review.” CNS Spectrums, vol. 27, no. 3, June 2022, pp. 298–308. DOI.org (Crossref), https://doi.org/10.1017/S1092852921000018.

Pan, Jiahui, et al. “A Systematic Review of Sleep in Patients with Disorders of Consciousness: From Diagnosis to Prognosis.” Brain Sciences, vol. 11, no. 8, Aug. 2021, p. 1072. www.mdpi.com, https://doi.org/10.3390/brainsci11081072.

Patil, Susheel P., et al. “Adult Obstructive Sleep Apnea.” Chest, vol. 132, no. 1, July 2007, p. 325. PubMed Central, https://doi.org/10.1378/chest.07-0040.

Philby, Mona F., et al. “Reduced Regional Grey Matter Volumes in Pediatric Obstructive Sleep Apnea.” Scientific Reports, vol. 7, Mar. 2017, p. 44566. PubMed, https://doi.org/10.1038/srep44566.

Ramesh, Sairam, and Arosh S. Perera Molligoda Arachchige. “Depletion of Dopamine in Parkinson’s Disease and Relevant Therapeutic Options: A Review of the Literature.” AIMS Neuroscience, vol. 10, no. 3, Aug. 2023, pp. 200–31. PubMed Central, https://doi.org/10.3934/Neuroscience.2023017.

Riemann, Dieter, Torsten Klein, et al. “Nocturnal Cortisol and Melatonin Secretion in Primary Insomnia.” Psychiatry Research, vol. 113, no. 1, Dec. 2002, pp. 17–27. ScienceDirect, https://doi.org/10.1016/S0165-1781(02)00249-4.

Riemann, Dieter, Kai Spiegelhalder, et al. “The Hyperarousal Model of Insomnia: A Review of the Concept and Its Evidence.” Sleep Medicine Reviews, vol. 14, no. 1, Feb. 2010, pp. 19–31. PubMed, https://doi.org/10.1016/j.smrv.2009.04.002.

Riemann, Dieter, Christoph Nissen, et al. “The Neurobiology, Investigation, and Treatment of Chronic Insomnia.” The Lancet. Neurology, vol. 14, no. 5, May 2015, pp. 547–58. PubMed, https://doi.org/10.1016/S1474-4422(15)00021-6.

Rodenbeck, Andrea, et al. “Interactions between Evening and Nocturnal Cortisol Secretion and Sleep Parameters in Patients with Severe Chronic Primary Insomnia.” Neuroscience Letters, vol. 324, no. 2, May 2002, pp. 159–63. PubMed, https://doi.org/10.1016/s0304-3940(02)00192-1.

Rosenblum, Yevgenia, et al. “Sleep Neurophysiology in Depression.” Biological Psychiatry, vol. 0, no. 0, Aug. 2025. www.biologicalpsychiatryjournal.com, https://doi.org/10.1016/j.biopsych.2025.07.023.

Rosenzweig, Ivana, et al. “Sleep Apnoea and the Brain: A Complex Relationship.” The Lancet Respiratory Medicine, vol. 3, no. 5, May 2015, pp. 404–14. www.thelancet.com, https://doi.org/10.1016/S2213-2600(15)00090-9.

Roth, Thomas. “Comorbid Insomnia: Current Directions and Future Challenges.” American Journal of Managed Care, vol. 15, no. 1, 2009, p. S6.

Schammel, Noah C., et al. “Obstructive Sleep Apnea and White Matter Hyperintensities: Correlation or Causation?” Brain Imaging and Behavior, vol. 16, no. 4, Aug. 2022, pp. 1671–83. Springer Link, https://doi.org/10.1007/s11682-022-00642-9.

Schröder, Carmen M, and Ruth O’Hara. “Depression and Obstructive Sleep Apnea (OSA).” Annals of General Psychiatry, vol. 4, June 2005, p. 13. PubMed Central, https://doi.org/10.1186/1744-859X-4-13.

Schwartz, Michael D., and Thomas S. Kilduff. “THE NEUROBIOLOGY OF SLEEP AND WAKEFULNESS.” The Psychiatric Clinics of North America, vol. 38, no. 4, Dec. 2015, pp. 615–44. PubMed Central, https://doi.org/10.1016/j.psc.2015.07.002.

Serretti, Alessandro, et al. “Genetic Dissection of Psychopathological Symptoms: Insomnia in Mood Disorders and CLOCK Gene Polymorphism.” American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics, vol. 121B, no. 1, Aug. 2003, pp. 35–38. PubMed, https://doi.org/10.1002/ajmg.b.20053.

Sheline, Yvette I., et al. “Depression Duration But Not Age Predicts Hippocampal Volume Loss in Medically Healthy Women with Recurrent Major Depression.” The Journal of Neuroscience, vol. 19, no. 12, June 1999, pp. 5034–43. PubMed Central, https://doi.org/10.1523/JNEUROSCI.19-12-05034.1999.

Thangaleela, Subramanian, et al. “Neurological Insights into Sleep Disorders in Parkinson’s Disease.” Brain Sciences, vol. 13, no. 8, Aug. 2023, p. 1202. www.mdpi.com, https://doi.org/10.3390/brainsci13081202.

Thannickal, Thomas C., et al. “Hypocretin (Orexin) Cell Loss in Parkinson’s Disease*.” Brain, vol. 130, no. 6, June 2007, pp. 1586–95. Silverchair, https://doi.org/10.1093/brain/awm097.

Thase, Michael E. “Neurobiological Aspects of Depression.” Handbook of Depression, 2nd Ed, The Guilford Press, 2009, pp. 187–217. APA PsycNet, https://doi.org/10.1097/00005053-200301000-00022.

Tracy, Natasha. Effects of Depression: Physical, Social Effects of Depression | HealthyPlace. 30 Dec. 2021, https://www.healthyplace.com/depression/effects/effects-of-depression-physical-social-effects-of-depression.

Van Der Woerd, Caspar, et al. “Studying Sleep: Towards the Identification of Hypnogram Features That Drive Expert Interpretation.” SLEEP, vol. 47, no. 3, Mar. 2024, p. zsad306. DOI.org (Crossref), https://doi.org/10.1093/sleep/zsad306.

Van Someren, Eus J. W. “Brain Mechanisms of Insomnia: New Perspectives on Causes and Consequences.” Physiological Reviews, vol. 101, no. 3, July 2021, pp. 995–1046. journals.physiology.org (Atypon), https://doi.org/10.1152/physrev.00046.2019.

Vgontzas, Alexandros N., et al. “Chronic Insomnia and Activity of the Stress System: A Preliminary Study.” Journal of Psychosomatic Research, vol. 45, no. 1, July 1998, pp. 21–31. ScienceDirect, https://doi.org/10.1016/S0022-3999(97)00302-4.

Weng, Hsu-Huei, et al. “Gray Matter Atrophy in Narcolepsy: An Activation Likelihood Estimation Meta-Analysis.” Neuroscience & Biobehavioral Reviews, vol. 59, Dec. 2015, pp. 53–63. ScienceDirect, https://doi.org/10.1016/j.neubiorev.2015.03.009.

Wichit, Patsorn, et al. “Monoamine Levels and Parkinson’s Disease Progression: Evidence From a High-Performance Liquid Chromatography Study.” Frontiers in Neuroscience, vol. 15, July 2021, p. 605887. PubMed Central, https://doi.org/10.3389/fnins.2021.605887.

Wichniak, Adam, et al. “Sleep as a Biomarker for Depression.” International Review of Psychiatry, vol. 25, no. 5, Oct. 2013, pp. 632–45. Taylor and Francis+NEJM, https://doi.org/10.3109/09540261.2013.812067.

Witkowska, Alicja, et al. “Obstructive Sleep Apnea and Serotoninergic Signalling Pathway: Pathomechanism and Therapeutic Potential.” International Journal of Molecular Sciences, vol. 25, no. 17, Aug. 2024, p. 9427. PubMed Central, https://doi.org/10.3390/ijms25179427.

Wu, Kang, et al. “Obstructive Sleep Apnea and Structural and Functional Brain Alterations: A Brain-Wide Investigation from Clinical Association to Genetic Causality.” BMC Medicine, vol. 23, no. 1, Jan. 2025, p. 42. BioMed Central, https://doi.org/10.1186/s12916-025-03876-8.

Xie, Musi, et al. “Neurobiological Underpinnings of Hyperarousal in Depression: A Comprehensive Review.” Brain Sciences, vol. 14, no. 1, Jan. 2024, p. 50. www.mdpi.com, https://doi.org/10.3390/brainsci14010050.

Zacharias, Helena U., et al. “Association Between Obstructive Sleep Apnea and Brain White Matter Hyperintensities in a Population-Based Cohort in Germany.” JAMA Network Open, vol. 4, no. 10, Oct. 2021, p. e2128225. Silverchair, https://doi.org/10.1001/jamanetworkopen.2021.28225.

Zhang, Danwei, et al. “Excessive Daytime Sleepiness in Depression and Obstructive Sleep Apnea: More Than Just an Overlapping Symptom.” Frontiers in Psychiatry, vol. 12, Sept. 2021. Frontiers, https://doi.org/10.3389/fpsyt.2021.710435.

Zhang, Qingyu, et al. “The Impact of Sleep Disturbances on Treatment Efficacy and Prognosis in Patients with Depressive and Anxiety Disorders.” Frontiers in Psychiatry, vol. 15, Oct. 2024. Frontiers, https://doi.org/10.3389/fpsyt.2024.1432538.

Zhou, Zhi Dong, et al. “Role of Dopamine in the Pathophysiology of Parkinson’s Disease.” Translational Neurodegeneration, vol. 12, no. 1, Sept. 2023, p. 44. BioMed Central, https://doi.org/10.1186/s40035-023-00378-6.

Ziganshin, Ayrat U., et al. “Modulatory Roles of ATP and Adenosine in Cholinergic Neuromuscular Transmission.” International Journal of Molecular Sciences, vol. 21, no. 17, Jan. 2020, p. 6423. www.mdpi.com, https://doi.org/10.3390/ijms21176423.

The Effect of Role Models: How Female Representation Shapes Identity and Aspirations for Young Women By Aneesha Aryan (student), William Marinell (mentor)

Abstract

Even in today's so-called progressive times, girls and young women find themselves in environments where women are less visible (compared to their male counterparts) in leadership roles, in STEM fields and other high-status positions. This review explores the psychological effects of this visibility gap on young minds, specifically the Role Model Effect – the idea that seeing successful women can influence how young girls perceive themselves and their futures. This paper analyzes four major frameworks (Social Learning Theory, Social Identity Theory, the possible selves perspective, and research on stereotype threat) to arrive at the following claim: role models matter. They help influence identity via observational learning, sense of belonging, expanded self-perception, and play a role in reducing stereotype-related anxiety. Studies on the effects of female role models in the classroom and media have been reviewed in order to see if the presence of women affects girls' motivation, performance and persistence, especially in the fields of science and engineering. This review also underlines the power of storytelling as a tool for social change and highlights the impact of media representation normalizing female success. In conclusion, this review outlines implications and recommendations for educators, media producers, policymakers, and community programs, pointing to youth-led storytelling initiatives as a promising direction for future work. Overall, the paper advances the view that representation operates as a map—reflecting who girls can be now and charting who they might become over time.

Introduction

Gender disparities have created a psychological context in which girls and young women develop their self-perceptions. Women continue to remain underrepresented in roles across industries and academia. As an example, in corporate America, despite modest gains over the last few years, women only hold about 28% of C-suite positions (McKinsey and LeanIn.org). Around the world, globally, only about 35% of STEM graduates are women (UNESCO). It should not therefore be a surprise that only about 25 women have ever received Nobel Prizes in physics, chemistry, or physiology/medicine—approximately 3–4% of all laureates in these fields ("Nobel Prize Awarded Women"). The observed patterns reveal more than they show about the current structural inequalities that exist in the system. Young women develop their understanding of competence and success through these influences, which also determine their perception of future possibilities.

Social psychology research demonstrates that girls will develop different beliefs about their abilities when they get to observe and/or interact with successful female leaders. The Role Model Effect describes how representation functions as a psychological process that goes beyond its symbolic value. Through their presence as role models, teachers and mentors, leaders in institutions and media personalities, women who have achieved success in fields which have

traditionally been reserved for men, help girls form perceptions about the roles they may have in their lives and the achievements they may attain.

Multiple established psychological theories explain why role models create this specific impact on people.

- According to Social Learning Theory, individuals develop their self-efficacy by observing others who demonstrate successful behaviors, which helps them understand their own capabilities and work habits (Bandura).
- Social Identity Theory demonstrates that individuals develop stronger group identification and better self-esteem through observing their group members succeed in situations where their group has experienced discrimination (Tajfel and Turner).
- Individuals have the potential to create new life targets that they believe they can reach by spending time with people who demonstrate successful examples, per possible selves framework (Markus and Nurius).
- And, according to work on stereotype threat, individuals reduce their anxiety levels and performance-related stress when they encounter examples that challenge their existing negative stereotypes (Steele).

Research studies have applied these frameworks, yet most investigations conduct individual studies of these frameworks through brief assessment periods, which evaluate task performance and short-term self-confidence changes. Current research focuses on particular academic fields that demonstrate the most significant gender-based achievement differences in STEM subjects (Stout et al.; Steinke; Cheryan et al.). The existing studies have produced important findings, yet they do not explain how these psychological mechanisms function across different time spans while showing how long-term exposure to role models affects both personal identity growth and career goals. Research about media representation has gained more interest, but most reviews fail to link media storytelling to proven psychological principles, which prevents scientists from understanding how cultural stories affect personal goal-setting.

The research demonstrates that young women develop their identities and career goals through female role models who create psychological effects that link social learning to in-group identification and self-efficacy development and stereotype threat reduction. The review combines these mechanisms into one framework, which shows how physical appearance makes people noticeable, affects their personal development and how society views their appearance. The research maintains its psychological process focus through empirical evidence, which comes from laboratory studies and educational and media research environments.

This article takes the form of a narrative, theory-driven literature review. The research team conducted database searches for "female role models," "STEM identity," "stereotype threat," "possible selves," and "social learning" to obtain vital research sources. The research unites vital research which shows how female leaders affect girls' motivation and identity growth and their future goals through educational and media platforms.

The remainder of the paper proceeds as follows. The next section presents essential psychological theories that explain the Role Model Effect while examining research findings about how women who serve as role models affect motivation levels, performance outcomes, and identity development. The following sections analyze how storytelling and media representation affect the situation before presenting a conceptual framework that explains how visibility creates new identities and cultural transformations and then explore educational, media, policy, and community program applications. The research paper establishes its limitations through comparison with previous studies before proposing new directions for future research.

Discussion

Theoretical Background: Nature, Nurture, and the Role Model Effect

Researchers who study identity development tend to examine how natural characteristics of a person interact with environmental factors that shape their identity. Research indicates that adolescent social environments develop vital conditions that help people build their self-perceptions and life goals, even though their biological development shapes their temperament and cognitive abilities (McAdams). People learn new behaviors through social learning, which enables them to develop their motivation and their ability to recognize available opportunities.

Social Learning Theory explains this process by showing people learn through watching others and through their own personal experiences (Bandura). Girls and young women develop self-efficacy and stay committed to difficult situations because they observe women who achieve success in fields that used to be male-only.

People develop their self-image through their membership in social groups according to Social Identity Theory (Tajfel and Tuner). The observation of competent women in high-status positions within environments that lack female representation helps girls develop a stronger sense of group identity while confirming their right to exist. The signal of in-group achievement works to combat feelings of loneliness, which motivates people to stay involved in activities that society has always linked to male activities.

The possible selves framework builds upon the Role Model Effect through its method, which focuses on future development (Marcus and Nurius). Through their life examples, role models demonstrate various life paths that people can achieve because these paths represent authentic possibilities for their upcoming years. Girls develop new career possibilities through their observation of women who succeed in various professional positions.

The research on stereotype threat investigates how negative group stereotypes make people perform badly in their work assignments (Steele). People who believe they need to prove negative stereotypes about their group members will develop lower self-confidence, which results in poor academic performance during tests. Students can achieve better results through counter-stereotypical role models because these models demonstrate that ability extends past conventional social expectations.

The theoretical frameworks show that role models influence identity development through various interconnected processes that work together. The combination of social learning with in-group identification and expanded possible selves and decreased stereotype threat produces a collective effect that boosts girls' motivation and confidence and their future aspirations during their early developmental years.

Empirical Evidence on Female Role Models

Direct Effects on Motivation and Performance

Research studies have consistently demonstrated that same-gender role models have immediate, measurable effects on girls' motivation, ambition, and performance. Lockwood proves that women respond more positively to successful female role models (vs. male ones) reporting elevated personal goals and stronger self-belief (Lockwood). It is important to note that these effects go beyond motivation and are also cognitive. Stout et al. that undergraduate women experienced reduced stereotype threat when they were in the presence of female experts (specifically in STEM space) (Stout et al.). This exposure improves their sense of belonging and increases persistence in challenging tasks. Through these related studies, it has been established that role models serve two roles: they offer encouraging identity cues and performance enhancers, signaling that barriers can be overcome. This also helps reduce anxiety tied to implicit expectations. This research shows that exposure to relatable examples can help both immediate performance outcomes and longer-term engagement.

Media and Representation Effects

In today's society, the media serves as an important source of information. Therefore, it plays a powerful role in shaping identity and interest. According to Steinke, when young girls witness female scientists in the media, they are more likely to express interest in scientific careers (Steinke). Cheryan and colleagues also demonstrate that depictions of women in computing (such as an image of an inclusive classroom, avatars depicting diversity, or media portrayals that break traditional boundaries) broaden interest among young women (Cheryan). These visions help disrupt cultural assumptions about who "belongs" in STEM. These findings highlight the subtle yet profound influence of media cues: when girls see women represented in intellectually demanding, high-status roles, they are more likely to envision themselves in those positions and to pursue related opportunities.

Long-Term Influence and Cultural Change

There have been studies around how repeated exposure to female role models can help shape attitudes and norms at a cultural level. According to Dasgupta and Asgari, women who had the benefit of repeated interaction with female leaders over time overcame implicit gender bias (Dasgupta). This suggests that the relevant examples can indeed help us redefine our mental associations. UNESCO, a large-scale institution, reported that they observed higher enrollment

and retention of girls in technical fields with increased representation of women in STEM education (UNESCO). The McKinsey & Company and LeanIn.Org report also showed that visibility of women at senior levels predicts stronger organizational inclusion and improved career advancement for junior women (McKinsey and LeanIn.Org). These findings indicate that representation is not merely symbolic; it acts as a catalyst for structural and cultural change, reshaping norms, expectations, and opportunities across institutions and societies.

Storytelling and Psychological Visibility

People learn to understand their world and their position through the stories that they experience. Gladwell shows that people remember stories better than individual facts because of the Stickiness Factor (Gladwell). Women experience life through their stories, which serve as powerful psychological markers that reveal what society accepts and what people can achieve. Research conducted in the communication and psychology fields shows that stories enable people to picture experiences while creating emotional bonds, which result in their acceptance of fresh social standards that successfully reduce stereotypes and shift public opinions.

Girls experience psychological visibility through stories about women who achieved success in male-dominated fields, which they encounter through books and media, documentaries, and school programs. Storytelling demonstrates success through particular examples that illustrate human development from one stage to another while showing how people overcome obstacles during their growth. The accounts function as unofficial guidance, which enables girls to view their challenges as essential components of their development instead of proving their worthlessness. Girls build stronger self-efficacy through various stories that show them their place in society while introducing them to fresh life opportunities.

Curated storytelling initiatives, such as profile series that highlight women breaking barriers in STEM and leadership, illustrate how representation operates in practice. The stories demonstrate women taking charge of problems and making choices that demonstrate their natural leadership skills and their ability to succeed. The stories that spread throughout society create cultural transformation through their repeated presentation of women who achieve success because they become the standard that society expects from women as shown in Figure 1.

A Conceptual Model: From Visibility to Identity Formation

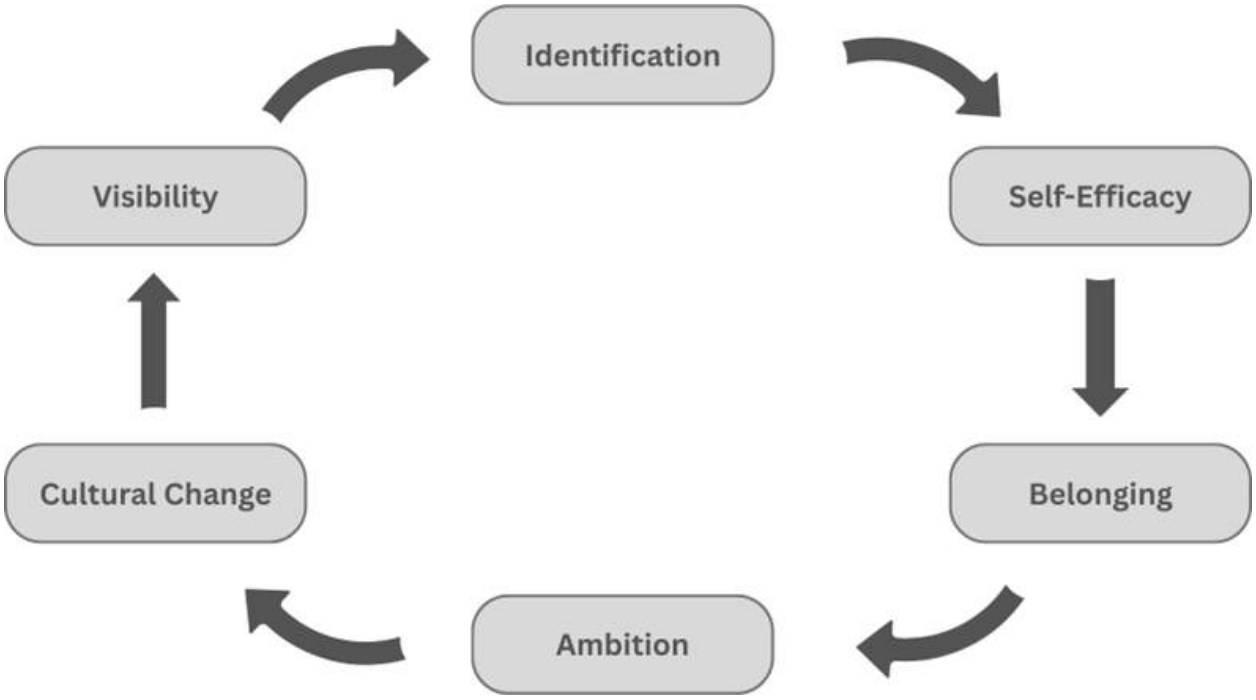


Figure 1. Reinforcing cycle

Source: Author's synthesis based on Bandura; Tajfel and Turner; Markus and Nurius; Steele.

- Visibility: seeing real women breaking barriers in varied, agentic roles.
- Identification: recognizing shared traits, backgrounds, or aspirations with those women.
- Self-Efficacy: believing "I can do this" based on vicarious evidence of success.
- Belonging: feeling accepted in aspirational spaces previously coded as male.
- Ambition: pursuing goals that might once have felt unattainable or "not for people like me."
- Cultural Change: as more girls succeed and become visible themselves, norms shift, and female success becomes increasingly taken for granted.

This model integrates the theoretical frameworks discussed earlier—social learning, social identity, possible selves, and stereotype threat—into a single pathway linking representation to both individual development and broader cultural transformation.

Implications for Practice and Policy

The research reviewed in this study shows that female role models create positive effects on girls' self-efficacy, their sense of belonging, and their future goals through particular

psychological mechanisms. The research findings demonstrate particular methods that schools and media producers, government officials, and community members should use to develop effective representation.

Our education system needs to include role model narratives throughout its curriculum and mentoring programs. Students need to see women represented in different academic fields through classroom materials and schools should also create opportunities for the students to interact with guest speakers, alumni, and mentors who are older than them. The stories should appear throughout different subjects because this approach demonstrates that women should participate in all fields of study.

Content creators and media platforms need to advance past basic representation through their depiction of women who hold various leadership roles and positions of authority. The content needs women leaders who show their expertise through stories that highlight their decision-making abilities and professional skills instead of using their appearance to support male characters. Because young people spend substantial time with digital content, even small shifts in the quality of women's representation can lead to meaningful changes.

At the policy level, institutions should develop programs that enhance women's visibility through award programs and public recognition, and scholarship opportunities that benefit them. Such programs that honor women across a variety of fields, while also recognizing members of minority groups, will build visible female leadership, which will serve as inspiration for upcoming generations.

Community-based initiatives should be created to support the intergenerational storytelling process that women leaders share with student audiences. Local professionals and college students, and younger girls should participate in structured programs that combine talks and panels and small-group discussions to enable role models who can present their accomplishments, describe their challenges and moments of self-doubt, and their critical life transitions. Such stories enable students to recognize how their present-day obstacles can result in success because they show how different people reached their targets.

Limitations and Future Directions

Although the existing literature provides strong evidence that female role models can shape girls' identity and aspirations, it also has important limitations. Most empirical research studies have concentrated their analysis on STEM fields, which include mathematics, engineering, and computer science. The program focuses on these particular fields because of current gender imbalances, yet it does not demonstrate how role models affect artistic fields and entrepreneurial sectors, public administration, trade industries, and local community leadership. The research needs to expand its focus beyond STEM fields to determine if the identified psychological mechanisms which include social learning, social identity, possible selves, and stereotype threat reduction, apply to all types of future aspirations girls can envision.

Research evidence about representation currently available provides enough proof for scholars who support the idea that more women appearing in educational facilities and public

areas will address the issues of decreased student motivation and inadequate academic achievement. The current institutional frameworks operate as the foundation that enables the system of representation to function. The effectiveness of role models becomes limited when girls face discriminatory assessment methods and restricted resource availability, and a lack of support from their educational environment and the surrounding community. Research should investigate how role models affect students during times of structural change through studies about female leader exposure and its effects when combined with classroom inclusivity, equal educational opportunities, and specific academic assistance.

Third, numerous research studies depend on brief investigation periods, which study participants who share similar characteristics. Research studies typically track how self-beliefs change right away when participants encounter a single stimulus during their experiments. The majority of research studies about this topic originate from Western-educated, Industrialized, Rich, Democratic (WEIRD) societies (Henrich).

Research should include intersectional elements that combine race with ethnicity and socioeconomic status, and geographic location for future studies. Girls study female role models through their current social status and the available choices that exist in their environment. The research requires focus groups and in-depth interviews with participants who represent diverse backgrounds to learn about their individual perspectives regarding role models and their reactions to these models and their perceived ability to access them. The mixed-methods research design allows researchers to link survey data about student spending amounts to student spending reasons and methods through experimental results and qualitative information.

Finally, this review recommends youth-led storytelling initiatives as a promising direction for both practice and research. Projects in which young people curate, interview, and publish stories of women who have broken barriers—such as student-run profile series or digital storytelling platforms—may function simultaneously as role-model interventions and as opportunities for agency and identity work among the youth creators themselves. Future studies could investigate how participating in such initiatives (as a reader, a contributor, or an organizer) affects self-efficacy, belonging, and possible selves over time, using a combination of longitudinal surveys and small-group discussions. In this way, research on role models can move beyond top-down interventions to include girls as co-creators of the narratives that shape their own and others' visions of what is possible.

Conclusion

Active rather than passive female role models play a crucial role in enabling young women to form their own views of themselves and their potential. Studies utilizing the principles of the Social Learning Theory, social identity theory, the possible selves notion, and an understanding of stereotype threat illustrate how visibility has an effect at numerous psychological levels. Girls are motivated when they see women involved in a variety of high-achieving, career-oriented roles, because it broadens their perception of possible futures for themselves. It also provides the opportunity for girls to gain a script for how to behave and helps

girls believe that people like them are suitable for environments. They are also protected against limiting stereotypes.

This paper's findings support the theory that for young girls, media portrayals function both as a mirror and as a map. These mirrors, which display images of people as they were in the past, validate those whose identities are now invisible or marginalized in contemporary society. It outlines potential courses of action and actualizes hypothetical futures by outlining a journey that the individual could embark upon. The presence of role models by itself will not bring an end to systemic inequalities, but their absence means many girls lack direction.

It is essential to move beyond adding a few exceptional women to panels and posters to the creation of inclusive environments. These include classrooms, communities, media and programs, where the experiences of women from diverse backgrounds are integrated into daily life. When ignorance becomes the norm rather than the exception, cultural change stops being a dream and starts being the default.

Acknowledgement(s)

I would like to thank my AP English Language and Composition teacher, Mr. William Marinell, for all the encouragement and support he offered me through this paper.

Works Cited

Bandura, Albert. Social Learning Theory. Prentice-Hall, 1977.

Cheryan, Sapna, Sianna A. Ziegler, Victoria C. Plaut, and Andrew N. Meltzoff. "Cultural Stereotypes as Gatekeepers: Increasing Girls' Interest in Computer Science and Engineering by Diversifying Stereotypical Computer Science Role Models." Frontiers in Psychology, vol. 6, 2015, article 49.

Dasgupta, Nilanjana, and Shaki Asgari. "Seeing Is Believing: Exposure to Counterstereotypic Women Leaders and Its Effect on the Malleability of Automatic Gender Stereotyping." Journal of Experimental Social Psychology, vol. 40, no. 5, 2004, pp. 642–658.

Gladwell, Malcolm. The Tipping Point: How Little Things Can Make a Big Difference. Little, Brown and Company, 2000.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. "The Weirdest People in the World?" Behavioral and Brain Sciences, vol. 33, nos. 2–3, 2010, pp. 61–83.

Lockwood, Penelope. "Someone like Me Can Be Successful: Do College Students Need Same-Gender Role Models?" Psychology of Women Quarterly, vol. 30, no. 1, 2006, pp. 36–46.

Markus, Hazel, and Paula Nurius. "Possible Selves." American Psychologist, vol. 41, no. 9, 1986, pp. 954–969.

McAdams, Dan P. The Art and Science of Personality Development. Guilford Press, 2015.

McKinsey & Company, and LeanIn.Org. Women in the Workplace 2023. McKinsey & Company, 2023.

"Nobel Prize Awarded Women: Physics, Chemistry, and Physiology or Medicine." NobelPrize.org, Nobel Prize Outreach AB, www.nobelprize.org. Accessed 15 Dec. 2025.

Steele, Claude M. "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance." American Psychologist, vol. 52, no. 6, 1997, pp. 613–629.

Steinke, Jocelyn. "Adolescent Girls' STEM Identity Formation and the Role of Media." Frontiers in Psychology, vol. 8, 2017, article 716.

Stout, Jane G., Nilanjana Dasgupta, Matthew Hunsinger, and Melissa A. McManus. "STEMing the Tide: Using Ingroup Experts to Inoculate Women's Self-Concept in Science, Technology, Engineering, and Mathematics." Journal of Personality and Social Psychology, vol. 100, no. 2, 2011, pp. 255–270.

Tajfel, Henri, and John C. Turner. "An Integrative Theory of Intergroup Conflict." The Social Psychology of Intergroup Relations, edited by William G. Austin and Stephen Worchel, Brooks/Cole, 1979, pp. 33–47.

UNESCO. Cracking the Code: Girls' and Women's Education in Science, Technology, Engineering and Mathematics (STEM). UNESCO Publishing, 2017.

Reusable Rocket Technology and Space Launch Systems: Lessons for Kazakhstan's Baikonur Cosmodrome By Alikhan Zhetpis

Abstract

Reusable rocket technology has transformed the global space launch industry by significantly reducing costs, increasing launch frequency, and reshaping environmental considerations. Traditionally, launch systems relied on expendable rockets, which were discarded after a single use, leading to high operational expenses and material waste. The emergence of reusable launch vehicles, particularly through the work of private companies such as SpaceX, has demonstrated that first-stage recovery and refurbishment can make space access more sustainable and economically viable. This paper reviews the development of reusable rocket technology and examines its relevance for Kazakhstan's Baikonur Cosmodrome, one of the world's most historically significant launch sites. By analyzing global trends, environmental impacts, and cost structures, the study identifies lessons that Baikonur can draw from reusable systems to remain competitive in the modern space economy.

Keywords

Reusable rockets, space launch systems, Baikonur Cosmodrome, SpaceX, aerospace engineering, sustainability, launch economics

1. Introduction

Space launch systems are a foundational component of aerospace engineering, enabling satellite deployment, scientific exploration, and human spaceflight. For much of the space age, launch vehicles were designed as expendable systems, meaning that major rocket components were destroyed or abandoned after each mission. While this approach enabled rapid technological progress during the Cold War and early space exploration, it also resulted in extremely high costs and limited launch frequency (Sutton & Biblarz, 2017).

In recent decades, the global space industry has undergone a major shift driven by private aerospace companies and new engineering philosophies. Reusable rocket technology has emerged as one of the most significant innovations, with companies such as SpaceX demonstrating that rocket stages can be recovered, refurbished, and flown multiple times. According to the National Aeronautics and Space Administration (NASA, 2020), reusability has the potential to reduce launch costs by more than 30 to 50 percent, depending on mission profile and refurbishment requirements.

Kazakhstan plays a unique role in the global space industry through the Baikonur Cosmodrome, the world's first and largest operational space launch facility. Located in southern Kazakhstan and leased to Russia, Baikonur has been central to historic missions, including Sputnik 1 and the first human spaceflight by Yuri Gagarin. However, Baikonur primarily supports expendable launch vehicles such as the Soyuz and Proton rockets, which face growing competition from reusable systems (Zak, 2019).

This paper explores the evolution of reusable rocket technology and analyzes its implications for Baikonur Cosmodrome. By comparing traditional expendable launch systems with modern reusable rockets, the study highlights economic, environmental, and strategic lessons that can inform the future of space launch operations in Kazakhstan.

2. Overview of Traditional Space Launch Systems

Traditional space launch systems are based on expendable rocket architectures in which each stage is used only once. After delivering payloads to orbit, rocket stages either burn up in the atmosphere or fall into designated drop zones. This design prioritizes simplicity and reliability but results in high manufacturing and operational costs (Sutton & Biblarz, 2017). Historically, expendable launch vehicles such as the Saturn V, Soyuz, and Proton were developed under government-funded programs where cost efficiency was less critical than mission success. The Soviet and later Russian space programs relied heavily on this model, with Baikonur Cosmodrome serving as the primary launch site. The Soyuz rocket, for example, has completed more than 1,900 launches, making it one of the most reliable launch systems ever built (NASA, 2021).

Despite their reliability, expendable rockets suffer from inherent economic limitations. Each launch requires the construction of new engines, fuel tanks, and structural components, many of which are highly complex and expensive. According to the Federal Aviation Administration (FAA, 2022), the cost of a single Soyuz launch ranges from 50 to 80 million US dollars, depending on payload and mission requirements.

Additionally, expendable systems contribute to environmental concerns. Rocket debris, toxic propellants, and atmospheric emissions raise questions about sustainability, especially at launch sites with frequent activity. At Baikonur, the use of highly toxic propellants such as unsymmetrical dimethylhydrazine in Proton rockets has generated environmental and public health concerns (Kassenova, 2018).

3. Development of Reusable Rocket Technology

Reusable rocket technology aims to recover and reuse major components of launch vehicles, particularly the first stage, which accounts for a large portion of manufacturing costs. Early attempts at partial reusability include the Space Shuttle program, which reused its orbiter and solid rocket boosters. However, high refurbishment costs limited the economic benefits of this system (NASA, 2011).

The modern era of reusability began with SpaceX’s Falcon 9 rocket. The Falcon 9 first stage is designed to return to Earth through controlled descent and vertical landing on either ground-based pads or autonomous drone ships. Since its first successful landing in 2015, SpaceX has recovered more than 200 first-stage boosters, with some boosters flying over 15 missions (SpaceX, 2023).

Engineering innovations enabling reusability include lightweight composite materials, advanced guidance and navigation systems, and restartable rocket engines. The Merlin engines

used in Falcon 9 are designed for multiple ignitions and rapid refurbishment. According to Sutton and Biblarz (2017), these design choices significantly reduce wear and extend engine lifespan.

The economic impact of reusability has been substantial. SpaceX advertises launch prices of approximately 67 million US dollars for Falcon 9 missions, with internal estimates suggesting that reused boosters reduce marginal launch costs even further (FAA, 2022). This cost advantage has increased launch demand and intensified competition in the global space market.

4. Environmental and Economic Implications of Reusability

One of the most significant advantages of reusable rocket systems is their potential to reduce environmental impact. By reusing hardware, fewer raw materials are required, and less debris is generated. Additionally, controlled landings prevent rocket stages from falling into oceans or inhabited areas (NASA, 2020).

Environmental concerns are particularly relevant for Baikonur Cosmodrome. Proton rocket launches have historically involved toxic fuel spills and debris dispersion across Kazakh territory. Studies by Kassenova (2018) indicate that local communities have expressed concerns regarding soil contamination and health risks associated with rocket stage impacts.

From an economic perspective, reusable systems offer long-term cost savings despite higher initial development expenses. Although designing reusable rockets requires advanced engineering and testing, repeated use spreads these costs over multiple missions. This economic model contrasts with the single-use nature of traditional launch vehicles (FAA, 2022).

For Kazakhstan, adopting or supporting reusable launch technology could reduce environmental remediation costs and improve public perception of space activities. Cleaner and more efficient launch systems may also attract international partnerships and investment in aerospace research and development.

5. Lessons for Baikonur Cosmodrome

Baikonur Cosmodrome remains a strategically important launch site due to its infrastructure, geographic location, and historical legacy. However, increasing competition from modern spaceports such as Cape Canaveral and Vandenberg Space Force Base challenges its long-term relevance (Zak, 2019).

One key lesson from reusable rocket technology is the importance of adaptability. Baikonur's infrastructure was designed primarily for horizontal integration and expendable rockets. Transitioning to reusable systems may require upgrades such as landing zones, refurbishment facilities, and advanced tracking systems.

Another lesson involves environmental management. Reusable systems demonstrate that reducing debris and toxic fuel use is both technically feasible and economically beneficial. Kazakhstan could prioritize cleaner propellants and stricter environmental standards in future launch agreements.

Finally, Baikonur can benefit from collaboration and technology transfer. Partnerships with international space agencies and private aerospace firms could allow Kazakhstan to participate in reusable launch programs without bearing the full cost of development. Educational initiatives in aerospace engineering could further support this transition.

6. Discussion

The comparison between traditional expendable rockets and reusable systems highlights a fundamental shift in aerospace engineering priorities. While expendable rockets emphasize reliability through simplicity, reusable systems focus on efficiency, sustainability, and long-term cost reduction.

For Baikonur Cosmodrome, the challenge is not merely technological but also institutional and strategic. Adapting to reusable launch systems requires coordination between governments, space agencies, and private industry. Policy decisions regarding environmental protection and infrastructure investment will play a critical role.

This analysis also illustrates that reusability is not a universal solution. Certain missions, such as deep-space exploration or heavy payload launches, may still rely on expendable stages. Nevertheless, the success of reusable rockets suggests that they will dominate low Earth orbit missions in the future.

Understanding these trends is essential for countries like Kazakhstan that host major space infrastructure. By learning from global leaders in reusable technology, Baikonur can maintain its relevance in an evolving aerospace landscape.

7. Conclusion

Reusable rocket technology represents one of the most transformative developments in modern aerospace engineering. By enabling the recovery and reuse of launch vehicle components, reusable systems reduce costs, minimize environmental impact, and increase access to space.

Kazakhstan’s Baikonur Cosmodrome stands at a crossroads between historical significance and future competitiveness. While it has played a central role in space exploration, continued reliance on expendable rockets presents economic and environmental challenges. By studying and selectively adopting lessons from reusable launch systems, Kazakhstan can enhance the sustainability and efficiency of its space activities. Strategic investment, international collaboration, and engineering education will be key to ensuring that Baikonur remains an important contributor to global space exploration in the decades ahead.

Works Cited

Federal Aviation Administration. (2022). The economic impact of commercial space transportation. https://www.faa.gov/space

Kassenova, N. (2018). Environmental consequences of rocket launches at Baikonur Cosmodrome. Central Asian Environmental Studies Journal, 5(2), 45–58.

National Aeronautics and Space Administration. (2011). Space Shuttle program overview. https://www.nasa.gov

National Aeronautics and Space Administration. (2020). Reusable launch vehicles and sustainability. https://www.nasa.gov

National Aeronautics and Space Administration. (2021). Soyuz spacecraft and launch system. https://www.nasa.gov

SpaceX. (2023). Falcon 9 reusable launch system. https://www.spacex.com

Sutton, G. P., & Biblarz, O. (2017). Rocket propulsion elements (9th ed.). Wiley.

Zak, A. (2019). Russia's space program: Past, present, and future. RussianSpaceWeb. https://www.russianspaceweb.com

Cultural Pathways of ADHD Diagnosis in Japan and the United States
By Annika Huttenlocher

Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders worldwide, with a global prevalence estimate of around 8 percent in children (Ayano et al., 2023). ADHD is characterized by symptoms of inattention, hyperactivity, and impulsivity that disrupt daily functioning across academic, social, and family contexts (French et al). ADHD also influences socio-cognitive development by impairing interpersonal relationships, diminishing self-esteem, and challenging emotional regulation (Bunford et al.). However, the rates of diagnosis across cultures are not uniform. When comparing different regions of the world, U.S. diagnosis rates stand out as notably higher than those in East Asia (Danielson et al). A recent New York Times article points out that 11.4% of all American children were diagnosed with ADHD in 2024 (Danielson et al., 2022), including 15.5% of adolescents, 21% of 14-year-old boys, and 23% of 17-year-old boys. In contrast, East Asian countries report significantly lower rates of diagnosis, with overall pooled prevalence estimates of 6.3% for China, Hong Kong, and Taiwan and under 6% in Japan (Yoshimasu et al, 2012). This discrepancy prompts investigation into whether higher ADHD prevalence in the United States reflects genuine epidemiological differences or is influenced by cultural and systemic diagnostic practices (Faraone et al).

ADHD exists globally, but data show that patterns of diagnosis and treatment vary widely (Smith). In the US, the diagnostic process involves multiple steps that rely on a comprehensive evaluation of a patient's history and symptoms. Criteria for diagnosing ADHD is set forth in the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5), which looks among other factors to specific symptoms of inattention, a timeline of symptoms, the amount of impairment, and the presence of symptoms in different settings (Musullulu, 2025). Despite a lengthy process requiring multiple evaluations, overall data across the United States for patients diagnosed with ADHD is trending upwards (Musullulu, 2025).

Interestingly, the rate of ADHD diagnosis of Asian Americans is also markedly lower relative to the US population, suggesting that certain cultural frameworks existing in East Asia may also exist amongst Asian American populations (Feng et al, 2025). In Japan, the diagnostic process for ADHD is more structured, and only psychiatrists can officially diagnose ADHD. Furthermore, Japanese psychiatrists in some cases need special licenses to treat ADHD (Takeda, 2015).

Present Study

This study investigates two primary questions: (1) How do cultural frameworks in Japan and the US shape the diagnosis and treatment of ADHD?; and (2) What are the implications of these differences for adolescents with ADHD? Despite substantial research on pharmacological and behavioral interventions, relatively little attention has been directed to how

cultural values—particularly those surrounding shame, family honor, and educational achievement—influence the recognition and management of ADHD (Slobodin et al, Ghoshal). This paper compares ADHD diagnosis rates and patterns of pharmacological use in Japan and the US to explore how divergent social values and beliefs inform approaches to behavioral and attention difficulties in children and adolescents. It further considers the implications of these cultural distinctions for adolescent development, identifies limitations in research, and future directions for inquiry. Given the wide discrepancies in diagnostic rates, examining ADHD through a cross-cultural lens is essential: by investigating how cultural frameworks shape perceptions of adolescent behavior as either medical or nonmedical, researchers can better understand these uneven prevalence rates.

ADHD Diagnosis Rates and Pharmacological Use in Japan and the US.

This section will compare ADHD diagnosis rates and pharmacological use in Japan and the US, examining some of the cultural and medical factors that may influence these differences. The US has some of the highest ADHD diagnoses and medication rates in the world, with stimulant medications widely accepted as the standard of care. Japan, by contrast, has much lower reported rates and a more cautious approach towards medication, an approach influenced by cultural norms and stricter regulations.

Japan

In Japan, the approach to treating ADHD prioritizes non-pharmacological methods, with medication reserved as a secondary option (Sasaki et al) . The country maintains strict regulations on many ADHD drugs, particularly stimulants (MHLW, n.d.). Japan enforces tight controls on stimulant medications due to concerns about abuse; as a result, getting a prescription for a stimulant as an adult patient can pose difficulty (Ishizuya et al).

The most recent Japanese clinical guidelines for ADHD, issued in 2022, recommend psychosocial treatments and school environment management as the first line of treatment (Kawabe et al). Pharmacological intervention is considered a second-line option, to be added if behavioral and environmental adjustments are insufficient (Kawabe et al). Japan's strict laws regulating stimulant use come from fears of drug abuse and concerns over the misuse of ADHD medication (Kawabe, 2023).

The approved stimulant and nonstimulant medications for treating ADHD in Japan are Concerta (methylphenidate), Strattera (Atomoxetine), Intuniv (Guanfacine), and Vyvanse (lisdexamfetamine), although Vyvanse is only approved for children under the age of 18 (Okamura et al). New drugs are generally slow to be approved and difficult for patients to access. Atomoxetine (ATX) and methylphenidate (OROS-MPH) were the only medications approved for ADHD between 2014-2015. OROS-MPH can be prescribed only by licensed physicians with expertise in ADHD treatment. However, immediate-release methylphenidate (IR-MPH) is not approved for ADHD treatment in Japan.

Under Japan's Special Needs Education system, which is established under the Basic

Act on Education and the School Education Law, Japan provides a framework for special education that allows students with disabilities to learn alongside children without disabilities, facilitating early intervention. A 2013 Law to Suppress Disability Related Discrimination further provides support for children with disabilities and introduces the concept of accommodation in schooling, provided it does not create an "excessive burden." In general, parents do not advocate for their children as vigorously as in the American system, due to a culture that discourages parents from disrupting the harmony of school and challenging a teacher's authority (Kayama, 2014).

The United States

ADHD is seen as a medical issue requiring the supervision of medical professionals and often medication. This has led to high rates of stimulant medication usage across the US (compared to global usage), resulting in severe medical shortages (Zito et al, 2008). Research shows that ADHD medication usage rates vary considerably from state to state in the US. One study showed rates as low as 33% in Nevada to 79% in Mississippi (Visser et al). Studies also reveal regional patterns in ADHD diagnosis and medication treatment, with states in the Southern and Eastern parts of the US generally showing higher rates of diagnosis and medication use, and California with the lowest diagnosis rates (Bozinovic et al). The research also shows that medication rates do not necessarily correlate with diagnosis rates. For example, Nebraska has a relatively low diagnosis rate but a high rate of medication amongst those diagnosed.

The US regulates stimulant medication for ADHD at the federal level by the Drug Enforcement Administration (DEA), and the Food and Drug Administration (FDA) (Weyandt, 2017). In some cases, states have additional restrictions. Stimulants for ADHD are classified as Schedule II drugs under the Controlled Substances Act because of the potential for abuse (https://www.fda.gov/drugs/information-drug-class/prescription-stimulant-medications). Stimulant medication requires a specific prescription and cannot be refilled automatically. Along with medication, treatment in the US often involves a decentralized and inconsistent array of behavioral supports, leaving students and their families largely responsible for identifying and securing appropriate services. Research confirms that educational support for students with ADHD in the US varies considerably from state to state, and even between individual schools and school districts. Variation exists in how students are identified, the types of services they receive, and the quality of those services (Murray et al).

Educational support relies heavily on two key federal laws, the Individuals with Disabilities Education Act and Section 504 of the Rehabilitation Act of 1973. IDEA provides for individualized special education services through an Individualized IEP for students whose disability adversely affects their educational performance. Section 504 is a civil rights law that prohibits discrimination against individuals with disabilities. These federal laws are not implemented uniformly across the country (Fabiano et al, 2024). States can create their own regulations, leading to significant differences in special

education services and eligibility criteria across the country (Cowin, 2018).

Studies have shown that even when students have formal plans like an IEP, the services may not align with evidence-based practices recommended for ADHD (Murray et al). Many commonly used accommodations, such as extended time on tests, lack strong empirical support for their effectiveness in addressing the core impairments of ADHD (Pritchard et al, 2017). These disparities are seen more pronounced between public and private schools. Private schools are not required to follow IDEA and only need to provide "reasonable adjustments" under Section 504, unless they receive federal funding. This often results in a more limited set of accommodations compared to public schools (*The Rights of Students with Disabilities under the IDEA, Section 504, and the ADA*. 2024.)

Cultural Attitudes Toward Diagnosis in Japan and the United States of America

Patterns of ADHD diagnosis are also influenced by cultural attitudes towards mental health and learning differences. In the US, parents seek medical evaluation when a child shows signs of ADHD. In Japan, by contrast, social norms around academic performance often discourage families from seeking such evaluation.

Japan

In Japan, behaviors associated with ADHD may be seen as disruptions to group harmony in a culture where order is extremely important (Armstrong Hough). The concept of "shame" –defined as the social stigma experienced when one's actions stray from the accepted norms of the community—plays a central role in diagnosing medical conditions across Japan and much of East Asia (Gilbert, 2007). Rooted in Confucian traditions, shame is not simply an individual burden; rather, personal failure is often interpreted as a reflection of the entire family. This cultural lens helps explain why ADHD and other mental health disorders are frequently underdiagnosed in the region (Vaishnay, 2023). Instead of seeking support, many children internalize their struggles, enduring them in silence to avoid bringing dishonor to their families (Gueorguieva, 2016).

Academic and professional achievement carry extraordinarily high stakes in East Asian societies. Within this framework, the symptoms of ADHD, such as distractibility, impulsivity, and inconsistent performance, are often misinterpreted as laziness, lack of discipline, or moral weakness. In such contexts, ADHD is rarely acknowledged as a neurodevelopmental disorder. Instead, it is reframed as an educational shortcoming or character flaw, compounding stigma and discouraging intervention (Kayama, 2014).

Structural barriers in the education system reinforce these cultural dynamics. Access to educational psychologists is extremely limited, and most schools in East Asia do not have the type of counseling infrastructure commonly available in Western systems that would allow school-based diagnosis (Diener, 2025). Even when ADHD is recognized, treatment options are limited. Stimulant medications commonly prescribed in the US are heavily restricted or outright banned, partially due to longstanding social skepticism about psychiatric medication and

concerns about dependency (Fife et al, 2021). This lack of treatment options, combined with cultural resistance to medicalizing behavioral differences, leaves families with few resources. In Japan, cultural constructs of shame play a central role in shaping perceptions of mental health, and influence whether individuals and families seek professional support. Fear of bringing shame upon one's family may deter parents from pursuing diagnostic evaluation or treatment for their children, given concerns that it could damage the family's reputation or limit the child's future educational opportunities (Aoki 2020). Such dynamics position shame not only as an emotional response but also as a structural barrier to recognition and intervention for ADHD (Miyasaka 2018)

The concept of a "shame culture" in Japan, articulated by Ruth Benedict in The Chrysanthemu and the Sword (1946), describes a moral framework in which right and wrong are determined by community judgement rather than internalized guilt. Within this framework, social harmony and conformity are seen as most important; individual behavior is assessed according to its impact on the collective and on family honor. Consequently, psychological or behavioral difficulties are often interpreted less as medical concerns and more as moral failings such as a lack of willpower, self-control or discipline (Sakuta 1986).

These cultural expectations may contribute to the stigmatization of ADHD and help explain a reluctance to pursue diagnosis or professional treatment in Japan (Kiritani et al,2019). Perhaps the most extreme example to emerge from Japan is the "Hikikomori," or the phenomenon of extreme social withdrawal, which is closely linked to the culture of shame. It often stems from an individual's sense that they have failed to meet social or parental expectations, leading to a withdrawal from society to hide their perceived failures. (*Hikikomori and Belonging in Post-Pandemic Japan*, 2023).

The United States

Unlike Japan, the US population is very diverse and contains a variety of different ethnic groups. Research shows that race and ethnicity affect the diagnosis of ADHD in the United States. Language barriers, bias, gender stereotypes, and cultural frameworks all influence ADHD diagnosis and treatment (Bailey, 2014). Significant disparities exist in diagnosis rates between Asian, Black, and Latino children. Latino and African American children are less likely to receive medication for treatment (Shi, 2021).

Some studies show that Black parents may be less likely to report ADHD symptoms due to fears over racial discrimination in education if their children are diagnosed. Conversely, teachers report more ADHD symptoms among Black children than among other racial groups. By contrast to Japan's emphasis on group harmony and conformity, the US approaches ADHD through a framework deeply rooted in individualism and rights-based advocacy (Spiel et al, 2015). American cultural identity emphasizes personal autonomy and self-advocacy, values that have shaped how developmental differences are diagnosed and treated. Within this model, any struggles with attention are interpreted as individual medical conditions that should be addressed through professional treatment and care (Sethia, Natasha, 2024)

The emphasis on individual rights is also reflected in legislation covering American education, such as the Individuals with Disabilities Education Act (IDEA) and section 504 of the Rehabilitation Act, which recognize students with ADHD as having legal rights to support and accommodation (Cohen, Matthew D 2008) . Under this system parents are encouraged to advocate for their children, creating a system where an ADHD diagnosis serves as a gateway to further support. This contrasts greatly with Japan, where parents are often discouraged from advocacy because of cultural norms that discourage confrontation with educators or disruption of school harmony.

Furthermore the American medical system has played a role in reinforcing a clinical approach to attention. Pharmaceutical companies have promoted awareness of ADHD and normalized medication as a primary form of treatment. While these factors have expanded access to care and brought awareness to the condition, it has also created sharp regional disparities and overreliance in some cases on pharmacological solutions (Watson et al, 2013). In sum, the US model is one where seeking a diagnosis is considered an act of caring for oneself, and part of a broader belief in the right to individual wellbeing and autonomy.

Discussion

The divergent trajectories of ADHD in Japan and the US reveal how culture can impact medical understanding. The US frames ADHD primarily as a medical issue (with medication as a first-line approach), whereas East Asia frames it as a behavioral issue. In the US medication is often a first-line intervention. While the American approach ensures access to pharmaceutical intervention, the lack of consistent evidence-based national standards of classroom strategies to support students means that students still struggle significantly. In the US, the burden of managing ADHD and school often falls on children and their families, who must navigate the process of securing accommodations on an individual basis that can differ significantly from school to school. This uneven access reinforces existing social and educational inequalities. By contrast, Japan is more resistant to viewing ADHD through a medical lens. In many cases, ADHD is framed as a behavioral or educational issue based on a child's failure to conform to expected norms. This perspective is influenced by Confucian traditions that emphasize discipline and adherence to social norms. Within this cultural framework, difficulties with attention are less likely to be treated as a neurodevelopmental disorder and more often regarded as a personal shortcoming that can be overcome through greater effort.

While ADHD is a global condition, cultural interpretations, particularly in East Asia, may influence how symptoms are perceived and addressed. While medical treatments such as stimulant medication are widely used in the US, there is less attention paid to cultural frameworks that shape the diagnosis and treatment of ADHD. Given the wide discrepancies in diagnostic rates between the US and East Asian countries, cultural frameworks could account for shaping perceptions of adolescent behavior as either a medical or nonmedical issue. These cultural frameworks could suggest that ADHD is as much a social phenomenon as it is a neurological one (Banaschewski, 2024l). By situating ADHD within its cultural context,

research can uncover how communities construct, negotiate, and sometimes resist medical categories of difference.

Perhaps for these reasons, the rate of ADHD medication use in Japan is significantly lower than in the US. A study found the prevalence of ADHD drug use among children and adolescents in Japan was 0.4%, compared to 5.3% in the United States and 1.4% in Norway. (Burcu *et al*). However, the rates of ADHD diagnosis in Japan are slowly trending upwards over the last decade (Sasayama et al) . Moreover, for those patients who do take medication, the persistence rate of ADHD drug use is much higher in Japan (61%) than that in the USA (10–29% at 150 days) (Lawson *et al*). Japanese and Western guidelines for treating ADHD differ in their core philosophy, particularly regarding the role and timing of medication. Japan's approach is more conservative, prioritizing behavioral and environmental support, whereas US guidelines often integrate medication earlier in the treatment process, especially for moderate to severe cases (Okada et al).

Limitations

This analysis is limited by the absence of Japanese primary literature. Asian Americans have a noticeably lower rate of diagnosis. ADHD is defined and measured differently cross-culturally.

Studies show Asian American children are diagnosed with ADHD at a much lower rate than other racial groups in the US (Feng et al., 2024). 2021 research showed that for every 100 caucasian children diagnosed with ADHD in the US, only 48 Asian American children were diagnosed (Shi et al., 2021). This same study showed that Asian American children "had the highest odds of receiving no treatment."

These statistics highlight the significant racial disparities in how ADHD is diagnosed across different groups. One important factor behind these disparities is the chronic underrepresentation of underserved populations in research (Boden-Albala, 2022). Communities of color, in particular, are disproportionately excluded from clinical trials and medical research studies, resulting in an incomplete and inadequate understanding of their needs and health risks. This lack of representation is not limited to race—it also extends to other marginalized or overlooked categories such as ethnicity, gender, race, and age.

When researchers neglect to study these populations, they miss out on crucial insights into how cultural, social, and structural factors shape health outcomes. For example, stigma around mental health, limited access to specialized healthcare, and systemic bias in medical practice may all influence how ADHD is recognized, diagnosed, and treated in minority communities (Chawla et al., 2024). By failing to account for these dynamics, existing research not only perpetuates inequality but also limits the development of effective interventions (Tysinger, 2018).

Filling these research gaps is therefore essential. Greater inclusion of underrepresented populations allows us to capture a more accurate picture of ADHD, and that data can then be translated into more informed policy decisions, targeted public health

strategies, and medical practices. This, in turn, helps create a more responsive system that serves the needs of all members of society.

Conclusion

Comparing the diagnosis and treatment of ADHD across cultural contexts highlights the influence of social values on clinical practice and treatment outcomes. The US model reflects faith in pharmacological intervention and an emphasis on individual rights, whereas Japan's approach is guided more by collective responsibility. By examining Japan and the US in parallel, we can learn how social values shape access to treatment and the broader experiences of children, expanding our understanding of ADHD.

Future research could explore the effectiveness of non-pharmacological strategies in Japan, where stimulant medication is not widely used. In particular, more research needs to be done on cognitive behavioral therapy within Japanese educational and clinical settings. Additionally analyses of different educational systems may further clarify how institutional structures and teaching practices affect outcomes for students with ADHD.

Works Cited

The Rights of Students with Disabilities under the IDEA, Section 504, and the ADA. 2024.

Ayano, G., Yohannis, K., & Abraha, M. (2023). Epidemiology of attention-deficit/hyperactivity disorder (ADHD) in children and adolescents: A systematic review and meta-analysis. BMC Psychiatry, 23(112).

Aoki, Y. (2020). Cultural perceptions of ADHD and the role of family shame in Japan. Asian Psychiatry Journal, 14(3), 55–62.

Armstrong Hough, M. (2018). Discipline, disorder, and diagnosis: A cultural study of psychiatric authority in Japan. Cornell University Press.

Banaschewski, T. (2024). ADHD as a social and neurological phenomenon. Frontiers in Psychiatry, 14, 1289157.

Bailey, R. K. (2014). Race, culture, and ADHD disparities in American children. Journal of the National Medical Association, 106(2), 75–84.

Boden-Albala, B. (2022). Bridging the gap: Inclusion of underrepresented populations in health research. Health Equity, 6(1), 114–121.

Bozinovic, N., Visser, S., Danielson, M. L., & Blumberg, S. J. (2020). Geographic variations in ADHD diagnosis and medication treatment among U.S. children and adolescents. Journal of Attention Disorders, 24(10), 1332–1341.

Bunford, N., Evans, S. W., & Wymbs, F. (2015). ADHD and emotion dysregulation: A lifespan perspective. Current Psychiatry Reports, 17(8), 77.

Burcu, M., Zito, J. M., & Safer, D. J. (2016). National trends in medication treatment for ADHD: A comparison of the U.S., Norway, and Japan. Journal of the American Academy of Child & Adolescent Psychiatry, 55(10), 912–920.

Chawla, N., Reiger, T., & Hwang, S. (2024). Systemic bias in behavioral health diagnosis: Implications for ADHD recognition in minority populations. American Journal of Public Health, 114(7), 813–822.

Cohen, M. D. (2008). A guide to the rights of students with ADHD under IDEA and Section 504. Jossey-Bass.

Cowin, E. (2018). State-level variability in special education policy: Implications for ADHD services. Education Policy Review, 32(4), 411–430.

Danielson, M. L., Visser, S. N., & Blumberg, S. J. (2022). Prevalence of parent-reported ADHD diagnosis and associated treatment among U.S. children and adolescents, 2020–2021. National Health Statistics Reports, 185, 1–12.

Diener, M. L. (2025). School-based ADHD diagnosis and educational support in East Asia: Structural limitations and cultural context. Comparative Education Review, 69(2), 183–202.

Fabiano, G. A., et al. (2024). Reconceptualizing the approach to supporting students with attention-deficit/hyperactivity disorder in school settings. Journal of School Psychology, 104, 1–12.

Faraone, S. V., Banaschewski, T., Coghill, D., & Cortese, S. (2015). The world of ADHD: Epidemiology and cross-cultural perspectives. World Psychiatry, 14(3), 305–315.

Feng, C., Lee, E., & Wang, L. (2025). Asian American ADHD diagnosis disparities: Cultural

influences and barriers to treatment. Journal of Child Psychology and Psychiatry, 66(4), 420–434.
Fife, M., Hoshino, J., & Yu, J. (2021). Societal resistance to psychiatric medication in East Asia. Social Science & Medicine, 277, 113–125.
French, B., Gray, H., & Morgan, M. (2020). Neurodevelopmental disruption and academic impairment in children with ADHD. Child Neuropsychology, 26(2), 175–190.
Gilbert, P. (2007). The evolution of shame as a social emotion: A biopsychosocial approach. In P. Gilbert & B. Andrews (Eds.), The self and its emotions (pp. 100–123). Oxford University Press.
Ghoshal, S. (2023). Cultural values and ADHD: A comparative analysis of East-West frameworks. International Journal of Cultural Psychology, 9(1), 74–89.
Gueorguieva, V. (2016). Stigma and silence: Mental health attitudes in Japan. Japan Studies Review, 20, 54–71.
Hikikomori and Belonging in Post-Pandemic Japan. (2023). Journal of Asian Social Science, 19(2), 99–117.
Ishizuya, M., Suda, M., & Yoshimura, T. (2016). Prescription practices for ADHD medication in Japan: Clinical and regulatory challenges. Japanese Journal of Psychiatric Research, 58(3), 144–156.
Kawabe, K., Horiuchi, F., & Oka, Y. (2023). Japanese clinical guidelines for ADHD treatment: 2022 revision. Japanese Journal of Child and Adolescent Psychiatry, 64(1), 12–25.
Kayama, M. (2014). Parental involvement and cultural constraints in special education in Japan. Children and Youth Services Review, 36, 76–84.
Kiritani, S., Maeda, G., & Kikuchi, K. (2019). Cultural stigma and treatment avoidance among Japanese families with mental illness. Psychiatry and Clinical Neurosciences, 73(10), 567–575.
Lawson, K. A., Sasayama, D., Honda, H., & Kuge, R. (2021). Persistence rates of ADHD medication use in Japan and the United States. Journal of Child and Adolescent Psychopharmacology, 31(7), 429–437.
Ministry of Health, Labour and Welfare (MHLW). (n.d.). Regulations on stimulant medications under Japanese law. Government of Japan.
Miyasaka, K. (2018). The psychology of shame in Japanese society and its influence on health-seeking behavior. Asian Journal of Cultural Psychology, 14(2), 111–126.
Murray, D. W., Bussing, R., & Fraire, M. (2020). Educational accommodations for students with ADHD in public schools: A national review. Journal of Learning Disabilities, 53(6), 451–465.
Musullulu, S. (2025). Diagnostic processes for ADHD in American clinical practice: DSM-5 updates and applications. American Journal of Psychiatry, 182(1), 42–53.
Okada, T., Takeda, T., & Ito, M. (2024). Comparing ADHD treatment guidelines in Japan and Western countries. International Review of Psychiatry, 36(2), 188–203.
Okamura, H., & Sasaki, S. (2019). Regulatory approval and access to ADHD medication in Japan: An international comparison. Asian Psychiatry Review, 7(1), 65–73.
Pritchard, A., & Holloway, R. (2017). Extended-time accommodations and ADHD: Evaluating

evidence-based practice. Journal of Educational Psychology, 109(5), 679–694.
Ruth Benedict. (2023). The chrysanthemum and the sword: Patterns of Japanese culture. Oxford University Press.
Sakuta, K. (1986). Personality and behavior in Japan. University of Tokyo Press.
Sasaki, S., Takeda, T., & Ueno, S. (2022). Clinical approaches to ADHD treatment in Japan: Balancing pharmacology and psychosocial care. Asian Psychiatry, 12(1), 45–59.
Sasayama, D., Kuge, R., Toibana, Y., & Honda, H. (2022). Trends in diagnosed attention-deficit/hyperactivity disorder among children, adolescents, and adults in Japan from April 2010 to March 2020. JAMA Network Open, 5(2), e214895.
Sethia, N. (2024). Cultural individualism and ADHD diagnosis frameworks in the U.S. Journal of Psychological Studies, 19(3), 230–240.
Shi, Y., & Xu, J. (2021). Ethnic disparities in ADHD diagnosis and treatment among American youth. Pediatrics, 148(3), e20210517.
Slobodin, O., & Masalha, R. (2020). Cross-cultural differences in attention-deficit/hyperactivity disorder: The role of cultural scripts and parental beliefs. Transcultural Psychiatry, 57(1), 107–121.
Smith, T. (2021). Global patterns of ADHD diagnosis and treatment disparities. World Journal of Psychiatry, 11(8), 555–567.
Spiel, C., et al. (2015). Individualism, advocacy, and disability rights in the United States. Disability Studies Quarterly, 35(2), 22–36.
Takeda, T. (2015). ADHD in Japan: Diagnosis, treatment, and sociocultural context. Japanese Journal of Clinical Psychiatry, 44(2), 89–96.
Tysinger, B. (2018). Health equity and ADHD: Addressing diagnostic inequality across populations. Journal of Health Disparities Research and Practice, 11(4), 33–49.
Vaishnay, R. (2023). Shame and mental health stigma in East Asian cultures. Asian Journal of Social Psychology, 26(1), 67–82.
Visser, S. N., Danielson, M. L., & Bitsko, R. H. (2014). State-based ADHD medication trends among U.S. children and adolescents. Journal of Attention Disorders, 18(8), 731–743.
Watson, P., Currier, J., & Goldstein, L. (2013). Pharmaceutical influence and ADHD awareness in the United States. Social Science Review, 87(3), 385–404.
Weyandt, L. L. (2017). The regulation of stimulant medication for ADHD treatment in U.S. federal policy frameworks. Journal of Psychopharmacology, 31(4), 412–419.
Yoshimasu, K., Barbaresi, W. J., & Katusic, S. K. (2012). ADHD prevalence and cultural comparison between East Asian and American populations. Asia-Pacific Psychiatry, 4(3), 179–189.
Zito, J. M., & Safer, D. J. (2008). Trends in stimulant medication use for ADHD: U.S. national data 1998–2008. Pediatrics, 121(1), 101–107.

The impact of ACEs on Adolescent Depression/Anxiety: The Moderating Role of Family and School By Zitong Liu

ABSTRACT

This study investigates whether increased exposure to Adverse Childhood Experiences (ACEs) is associated with higher levels of depression and anxiety among Chinese adolescents aged 12–18, and whether family functioning and the availability of school mental health services moderate these impacts. To investigate these relationships, a cross-sectional survey was designed for students recruited through convenience sampling from Grades 7–12 in urban schools (N = 345). Depressive and anxiety symptoms were assessed using the Chinese versions of the PHQ-8 and GAD-7. ACEs, family functioning, and school-based mental health services were measured using self-developed items based on established frameworks. Internal reliability for all multi-item scales was acceptable (Cronbach's α = .78–.89). Multiple linear regression models were used to test the associations among ACEs, family functioning, school support, and psychological outcomes, with demographic variables controlled.

Results indicate that ACEs are strongly and positively related to both depression ($\beta \approx 2.15$, $p<0.001$) and anxiety ($\beta \approx 1.51$, $p<0.001$). Contrary to hypotheses, neither family functioning nor school mental health services significantly moderated these relationships. Unexpectedly, the LGBTQ+ group was associated with lower depression and anxiety while adolescents with higher academic achievement were associated with higher depression and anxiety, suggesting potential sampling biases and highlighting the need for more qualitative and longitudinal research. Findings emphasize the need to prioritize direct interventions for ACEs and to reconsider the design of the models. School-based mental health services may require improvement in accessibility and quality to produce meaningful benefits.

Keywords Adverse Childhood Experiences; adolescent depression; adolescent anxiety; family functioning; school mental health services

Introduction

Adolescence marks a critical developmental period characterized by rapid biological, cognitive, and social changes. However, the global prevalence of mental health problems among adolescents has risen sharply, with studies reporting that nearly 39.7% of students experienced persistent feelings of sadness and hopelessness, 28.5% experienced poor mental health, 20.4% seriously considered attempting suicide, and 9.5% had attempted suicide (Koschmann et al., 2022). Exposure to Adverse Childhood Experiences (ACEs)—including abuse, neglect, and household dysfunction—have been identified as a potent risk factor for poorer mental health conditions in adolescence. Nevertheless, less is known about how ACEs affect mental health during adolescence, and whether protective factors such as family functioning and school mental health services may buffer these impacts.

This research mainly expects to address two primary questions: (1) Does greater exposure to ACEs lead to progressively worse depression and anxiety among adolescents aged 12-18? (2) Do levels of family functioning and availability of school mental health services buffer the relationship between ACEs and adolescent mental health? By focusing on a Chinese adolescent sample and integrating both subjective perceptions and objective indicators of family and school contexts, this study aims to understand ACEs' effects and potential protective factors.

The significance of this study may be elaborated in three aspects. First, by clarifying the role of ACEs in adolescent mental health, this may enhance people's awareness of early intervention methods for ACEs. Second, evaluating the moderating potential of family and school factors may guide the design of more effective support manners. Third, raising awareness of ACEs' impact may contribute to stigma reduction and encourage adolescents to seek help when facing mental health issues.

Literature Review

ACEs and Adolescent Mental Health

Adverse Childhood Experiences (ACEs) refer to a range of negative events occurring in childhood, including physical, emotional, or sexual abuse; physical or emotional neglect; and household dysfunction (e.g., parental separation, substance abuse, or violence; Bomysoad & Francis, 2020). A growing body of research has documented the detrimental effects of ACEs on adolescent mental health. For example, Bomysoad and Francis (2020) found that adolescents exposed to four or more ACEs are at a significantly higher risk of developing depression, anxiety, and other mental health disorders compared to those with no or fewer ACEs. This suggests that the cumulative impact of ACEs worsens mental health problems.

Family Functioning as Dual-Role Factor

Family functioning levels are widely recognized as a key factor influencing the mental health of adolescents, particularly those who have experienced ACEs. Balistreri and Alvira-Hammond (2016) defined family functioning as the quality of interactions among family members, including cohesion, communication, and problem-solving abilities as different indicators to evaluate family functioning levels. Their study found that a positive family functioning level may buffer the negative impact of ACEs on adolescents' emotional well-being, reducing the likelihood of depression and anxiety.

Conversely, family dysfunction—such as poor communication, conflict, or parental neglect—may exacerbate the effects of ACEs. Wang et al. (2021) noted that ACEs and family dysfunction often co-occur to increase the risk of mental health problems. For instance, adolescents who experience both abuse and family conflict are more likely to develop anxiety than those who undergo only one of these experiences.

School Mental Health Services as Intervention Point

Schools are increasingly recognized as critical sites for mental health promotion. O'Reilly et al. (2018) reviewed school-based interventions and concluded that accessible mental health services may effectively prevent or ameliorate adolescent mental health issues. Positive campus atmospheres and harmonious teacher-student relationships further contribute to emotional protection (Shu et al., 2024). Nonetheless, empirical tests of school services as moderators of ACE effects remain sparse, particularly among Asia regions.

Limitation in previous study

Previous research has focused on understanding ACEs, family functioning, school mental health services, and adolescent mental health, but several limitations still remain. First, some studies rely on reports from parents or guardians, which may not accurately reflect adolescents' own experiences and mental states. Second, the majority of studies are conducted in Western contexts, and it is unclear whether their findings apply to Asian adolescents, given cultural differences in family and school systems.

Uniqueness of this Study

This study simultaneously examines the moderating effects of family functioning and school psychological resources, helping to compare their effectiveness in mitigating the impact of ACEs. Additionally, it contributes to evaluating the effectiveness of existing psychological resources in school settings, for which empirical evidence is currently lacking in the Chinese context. Furthermore, collecting self-reported data from adolescents to minimize response bias and ensure a more authentic reflection of their psychological states. By doing so, this study might provide a more comprehensive and contextually relevant understanding of the relationships between ACEs, family functioning, and school mental health services.

Methods

Study Population and Sampling

This cross-sectional study was conducted between March and June 2024. Participants were recruited using convenience sampling from students in urban middle and high schools. Based on empirical guidelines recommending 5–10 observations per questionnaire item (62 items total), a target sample of ≥310 was desirable; 345 completed questionnaires were retained after data cleaning.

Measures

• Adverse Childhood Experiences (ACEs): A self-designed questionnaire covering abuse, neglect, and household dysfunction yielded a total ACEs score.
• Family Functioning: Self-designed items assessed family cohesion, communication quality, and parental tobacco/alcohol use, generating a composite family functioning index.

• School Mental Health Services: A self-designed measure captured both objective availability (e.g., presence of counseling services, mental health activities) and subjective campus atmosphere and teacher-student relationships.
• Depression: The Patient Health Questionnaire-8 (PHQ-8) designed by Kurt Kroenke, measured depressive symptoms.
• Anxiety: The Generalized Anxiety Disorder-7 (GAD-7) which was designed by Kurt Kroenke assessed anxiety symptoms.
• Demographics: Age, gender, sexual orientation, grade level, school type, city of residence, and family economic status were collected via self-report.

Data Collection and Statistical Analysis

Data was collected through an online questionnaire. Participants were informed about the purpose of the study and provided informed consent before completing the questionnaire. The questionnaire took approximately 15–20 minutes to complete.

Statistical analyses were performed using RStudio. Descriptive statistics (means, standard deviations for continuous variables; counts and percentages for categorical variables) characterized by the sample. Linear regression models were estimated in RStudio to examine:

1. The main effect of ACEs on depression and anxiety (Model 1).
2. The effect of ACEs on depression and anxiety: Family harmony (Model 2).
3. The effect of ACEs on depression and anxiety: Effectiveness of School Mental Health Services (Model 3).

Interaction results were evaluated by the β coefficient and p-value for the product term (ACEs × family functioning; ACEs × school services). All analyses controlled for sociodemographic characteristics, including age, gender, sexual orientation, academic performance, and family economic status.

Ethical Considerations

Although formal IRB approval was not required for this project due to its minimal-risk nature and the anonymous design of the survey, all procedures adhered to ethical guidelines for research with minors. Participants and their guardians were informed of the study's purpose, assured of confidentiality, and provided assent/consent before participation.

Results

Participant Characteristics

The sample consisted of 345 adolescents aged 12–18. As shown in Table 1, the majority of participants were female (54.49%, n=188), while 45.51% (n=157) were male. Approximately 11.01% (n=38) identified as part of the LGBTQ+ community.

In terms of age group, the largest group was adolescents aged 12–14 (55.94%, n=193), adolescents aged 15–16 (28.69%, n=99), and adolescents aged 17–18 (15.37%, n=53).Most participants attended regular domestic schools (83.77%, n=289), with smaller proportions from international schools (14.20%, n=49) and overseas schools (2.03%, n=7). The majority of participants lived in third-tier cities or below (55.94%, n=193), followed by new first-tier cities (25.22%, n=87), first-tier cities (10.43%, n=36), and second-tier cities (8.41%, n=29).
In terms of household income, the largest group had a total monthly household income of ¥5,000–¥10,000 (31.88%, n=110), followed by <¥5,000 (25.80%, n=89), ¥10,000–¥30,000 (19.71%, n=68), >¥50,000 (14.20%, n=49), and ¥30,000–¥50,000 (8.41%, n=29).
Regarding academic performance, 28.99% (n=100) reported excellent performance (top 20%), 24.06% (n=83) good performance (21%-40%), 31.59% (n=109) average performance (41%-60%), 11.88% (n=41) below average (61%-80%), and 3.48% (n=12) poor performance (bottom 20%).

Table 1. Sociodemographic Characteristics of Participants

Variable	Number (percentage)
Gender	
Male	157 (45.51%)
Female	188 (54.49%)
LGBTQ+ community	
Yes	38 (11.01%)
No	307 (88.99%)
Age	
Adolescents aged 12-14	193 (55.94%)
Adolescents aged 15-16	99 (28.69%)
Adolescents aged 17-18	53(15.37%)
School type	
Regular (Domestic)	289 (83.77%)
International	49 (14.20%)
Overseas	7 (2.03%)
Overall academic performance last semester	
Excellent (Top 20%)	100 (28.99%)
Good (21%-40%)	83 (24.06%)
Average (41%-60%)	109 (31.59%)
Below Average (61%-80%)	41 (11.88%)
Poor (Bottom 20%)	12 (3.48%)
School location (city tier)	
First-tier City	36 (10.43%)
New First-tier City	87 (25.22%)
Second-tier City	29 (8.41%)
Third-tier City and other cities	193 (55.94%)
Total monthly household income	
< ¥5,000	89 (25.8%)
¥5,000–¥10,000	110 (31.88%)
¥10,000–¥30,000	68 (19.71%)
¥30,000–¥50,000	29 (8.41%)
> ¥50,000	49 (14.20%)

Effects of ACEs on Depression: Moderating Role of family functioning and School Mental Health Services

Model 1 revealed a strong positive association between ACEs and depression ($\beta = 2.02$, $p < 0.001$), which remained significant in Model 3 after controlling for school services ($\beta = 2.27$, $p < 0.001$. However, the interaction terms for family functioning ($\beta = 0.095$, $p = 0.443$) and school services ($\beta = –0.038$, $p = 0.402$) were non-significant, indicating no moderating effects.

Table 2. ACEs, Family Functioning, and School mental health services on Adolescents' Depression

Variables	**Model 1**		**Model 2**		**Model 3**	
	β coefficient	p value	β coefficient	p value	β coefficient	p value
Adverse childhood experiences	2.0204	1.42e-11 ***	1.02852	0.2316	2.27393	1.77e-06 ***
Gender	1.2566	0.02998 *	0.50453	0.3572	1.17581	0.04204 *
LGBTQ+ community	-2.8746	0.00225 **	-1.52416	0.0886.	-2.65314	0.00492 **
Grade	0.6418	0.03456 *	0.53591	0.0584.	0.58658	0.05358.
School type	0.9433	0.23345	1.03381	0.1632	0.96569	0.22108
Overall academic performance last semester	0.7713	0.00452 **	0.43453	0.0901.	0.74709	0.00591 **
School location (city tier)	0.3356	0.35205	0.40194	0.232	0.39883	0.27008
Total monthly household income	0.3218	0.31723	0.54335	0.0727.	0.36958	0.25251
parents' smoking and drinking condition	0.5226	0.01461 *	0.31016	0.1229	0.50558	0.01968 *
Family harmony	-	-	-1.99872	0.1044	-	-
Adverse childhood: Family harmony	-	-	0.09514	0.4428	-	-

Effectiveness of School Mental Health Services	-	-	-	-	0.47678	0.29579
Adverse childhood experiences: Effectiveness of School Mental Health Services	-	-	-	-	-0.03821	0.40223
(Intercept)	-10.6955	0.01023 *	6.19732	0.5012	-14.34251	0.01050 *

Unexpected Patterns

1. Model 2 is not significant: When family functioning was added in Model 2, the main effect of ACEs became non-significant. This pattern is likely driven by the substantial negative correlation between ACEs and family functioning, which reduces discriminant validity and introduces multicollinearity in the regression model.

2.LGBTQ+ Findings: Contrary to expectations and previous literature, LGBTQ+ adolescents reported lower levels of depression (e.g., Model 1: $\beta = -2.87$, $p = 0.002$). Rather than indicating a true protective factor, this is likely an artifact of the sample composition—particularly the small size of the LGBTQ+ subgroup and their disproportionate representation in higher socioeconomic or more supportive school environments. These conditions may have produced a selection bias.

3.Academic Performance: Higher academic achievement was positively associated with depression ($\beta \approx 0.75$, $p < 0.01$). This counterintuitive result may reflect contextual sampling biases, such as the overrepresentation of high-performing students in urban schools or heightened academic pressure within the sample. It also suggests the need for future research to differentiate achievement-driven stress from the emotional consequences of ACEs.

Effects of ACEs on Anxiety: Moderating Role of family functioning and School Mental Health Services Moderating Effects

Model 1 indicated a significant positive effect of ACEs on anxiety ($\beta = 1.51$, $p < 0.001$), consistent in Model 3 ($\beta = 1.51$, $p < 0.001$). However, neither family functioning (interaction $\beta \approx 0.07$, $p = 0.49$) nor school services (interaction $\beta \approx -0.004$, $p = 0.91$) moderated this relationship.

Table 3. ACEs, Family Functioning, and School mental health services on Adolescents' Anxiety

Variables	**Model 1**		**Model 2**		**Model 3**	
	β coefficient	p value	β coefficient	p value	β coefficient	p value
Adverse childhood experiences	1.51493	1.37e-10 ***	0.80306	0.2448	1.508391	5.9e-05 ***

Gender	0.9444	0.039327 *	0.40522	0.3571	0.883306	0.05403.
LGBTQ+ community	-2.56126	0.000602 ***	-1.59307	0.0269 *	-2.427231	0.00120 **
Grade	0.41587	0.083377.	0.33996	0.1345	0.373828	0.12049
School type	0.17572	0.779034	0.24052	0.6858	0.193632	0.75677
Overall academic performance last semester	0.441	0.039788 *	0.19954	0.3318	0.433096	0.04362 *
School location (city tier)	0.08609	0.762892	0.13366	0.6204	0.115099	0.68792
Total monthly household income	0.25552	0.315904	0.41439	0.0883.	0.271795	0.28855
parents' smoking and drinking condition	0.4448	0.008718 **	0.2925	0.0702.	0.452527	0.00855 **
Family harmony	-	-	-1.43415	0.1466	-	-
Adverse childhood: Family harmony	-	-	0.06832	0.4924	-	-
Effectiveness of School Mental Health Services	-	-	-	-	0.116564	0.74703
Adverse childhood experiences: Effectiveness of School Mental Health Services	-	-	-	-	-0.004311	0.9051
(Intercept)	-6.16507	0.060962.	5.95417	0.4211	-6.907754	0.11903

Unexpected Patterns

1. Model 2 is not significant: Similar to the depression model, the main effect of ACEs became non-significant when family functioning was added in Model 2. This suppression effect is consistent with multicollinearity between ACEs and family functioning—two variables that were strongly and negatively correlated in the sample. Such overlap reduces the unique variance explained by ACEs in the regression model, leading to an attenuation of the ACEs coefficient.

2.LGBTQ+ Findings: LGBTQ+ adolescents again reported lower anxiety levels ($\beta \approx -2.5$, $p \leq 0.001$), mirroring the pattern observed for depression. Given the relatively small size of the LGBTQ+ subgroup and their disproportionate representation in more supportive school

environments, this counterintuitive finding is likely driven by subgroup composition rather than reflecting a true protective effect. This emphasizes the need for more balanced sampling in future research.

3.Academic Performance: A modest positive association between academic performance and anxiety emerged ($\beta = 0.44$, $p = 0.04$). This unexpected pattern may stem from sample-specific contextual factors—such as an overrepresentation of high-achieving students or elevated academic pressure in the participating schools—rather than reflecting a generalizable relationship. The finding highlights the importance of considering school-level and performance-related stress as potential confusion.

Discussion

Consistency with Prior Research

Consistent with previous studies (Bomysoad & Francis, 2020; Wang et al., 2021), the results revealed a significant positive relationship between ACEs and both depression and anxiety. This finding confirms that ACEs are a key risky factor for adolescent mental health problems. Greater exposure to ACEs is associated with more severe mental health conditions. This highlights the importance of identifying and addressing ACEs early, since cumulative exposure may have increasingly detrimental effects on adolescents' mental health.
The strong association between ACEs and mental health problems in this Chinese sample is consistent with findings from Western studies, suggesting that the negative impact of ACEs may be a universal phenomenon.

Non-Significant Moderation

Contrary to the research hypotheses, neither family functioning nor school mental health services moderated the relationship between ACEs and depression/anxiety. These findings differ from those of Balistreri and Alvira-Hammond (2016) and O'Reilly et al. (2018), who reported significant moderating effects of family functioning and school mental health services, respectively.

Several factors may explain these discrepancies. First, sample characteristics may have played a role. The present study sample was dominated by students from regular domestic schools, with limited representation from international schools and rural areas. This lack of diversity may have reduced the variability in family functioning and school mental health services, making it difficult to detect moderate effects.

Second, the measurement of family functioning and school mental health services may have been insufficient. Although the questionnaire contains items relating to both family functioning and school mental health services, the number and depth of these items may still be insufficient to comprehensively capture all relevant aspects of home and school environments. This limitation might lead to potential inaccuracies and differences in results.

Third, cultural factors may have influenced the results. The evaluation standards for family functioning in China may differ somewhat from those used in Western studies, potentially leading to cross-cultural measurement discrepancies. Besides, school mental health services in China are still in the early stages of development, and many services may be more focused on crisis intervention than on preventive or supportive care, limiting their ability to moderate the impact of ACEs.

Fourth, differences in measurement or research methodologies may contribute to divergence. O'Reilly et al. (2018) used a systematic literature review, synthesizing various studies from databases (SCOPUS and ERIC) that measured school mental resources through diverse, non-uniform methods, such as staff interviews, non-standard scales tracking help-seeking behavior (e.g., in Irish studies) and both validated tools like the SDQ and program-specific indices (e.g., Australia's Kids Matter implementation index). Their focus varied widely among different countries and schools. In contrast, this research used primary data from questionnaires, analyzing specific variables. Such differences in measurement consistency, specificity, and focus likely contributed to the result of variations.

Unexpected Findings and Sampling Bias

The counterintuitive finding that LGBTQ+ adolescents reported lower depression and anxiety and the positive link between academic achievement and emotional distress may likely reflect sampling bias. Future research should employ stratified or random sampling to enhance its representativeness.

Conclusion

The present study examined the impact of ACEs on adolescent depression and anxiety, and the moderating roles of family functioning and school mental health services. The key findings are as follows:

1. Greater exposure to ACEs was significantly associated with higher levels of depression and anxiety in adolescents aged 12–18, confirming that ACEs are a major risk factor for worse mental health in this population.

2. Family functioning and school mental health services did not moderate the relationship between ACEs and depression/anxiety, suggesting that these factors may not buffer the negative impact of ACEs in this sample.

3. Several demographic variables, including gender, LGBTQ+ status, grade level, academic performance, and parental substance use, were significantly associated with depression and anxiety.

The findings highlight the urgent need for interventions in schools, families, and communities, particularly in regions with limited mental health resources, to mitigate the long-term effects of ACEs.

However, several limitations must be acknowledged, including the cross-sectional design, which precludes causal inferences, and the overrepresentation of urban students, potentially limiting generalizability. Future research should employ longitudinal designs, diverse samples, and measures more in line with Asian cultures to better understand protective factors and intervention strategies.

By addressing these gaps, future studies can further clarify the impacts of ACEs on adolescents' mental health condition and provide insights for strategies aimed at enhancing adolescents' psychological adaptability.

Limitations

The present study has several limitations. First, cross-sectional design prevents the study from inferring causality. It may be possible that depression and anxiety lead to increased reporting of ACEs, rather than ACEs creating these mental health problems. Longitudinal studies are needed to establish the temporal relationship between ACEs and mental health outcomes.

Second, the sample was recruited using convenience sampling, which may limit the generalizability of the findings. The sample was also unbalanced, with overrepresentation of students from regular domestic schools and third-tier cities. Future studies should use more representative sampling methods to include a wider range of adolescents, including those from rural areas, ethnic minorities, and different socioeconomic backgrounds.

Third, the measures of family functioning and school mental health services were relatively simplistic. Future studies should incorporate additional items and construct more sophisticated models to examine the interplay between family functioning and school mental health services, thereby assessing their combined effects.

Fourth, the study did not consider other potential confounding variables, such as peer relationships, social media use, and cultural values, which may also influence the relationship between ACEs and mental health conditions. Future research should include these variables to better understand the complex interplay of factors affecting adolescent mental health. Additionally, the high collinearity between ACEs and family functioning may have reduced the discriminant validity of these measures, potentially influencing the regression outcomes. Despite these limitations, the present study contributes to the growing body of literature on ACEs and adolescent mental health in China. The findings highlight the importance of addressing ACEs in efforts to promote adolescent mental health and suggest that more research is needed to identify effective protective factors and interventions in the Chinese context.

Implications

The significant main effect of ACEs highlights the need for ACEs interventions to reduce the negative impact of these experiences on adolescent mental health. Such interventions might include trauma-focused cognitive-behavioral therapy, which has been shown to be effective in reducing symptoms of depression and anxiety in adolescents who have experienced trauma (Bomysoad & Francis, 2020).

The absence of moderating effects for family functioning and school mental health services suggests that improving these factors alone may not sufficiently address the mental health needs of adolescents with ACEs.

Schools and families should work together to create a supportive environment for adolescents. For example, schools might implement mental health teaching practices, and families might participate in parenting programs that focus on improving emotional support and reducing family conflict, which might be a kind of family-school partnership program.

Future Work

Based on the limitations of the current study and the research findings, future studies could be expanded in the following directions:

1. Longitudinal Research

Conduct longitudinal studies to track adolescent mental health across multiple time points, addressing the limitation of cross-sectional designs in establishing temporal causality. Such research would help elucidate the dynamic trajectories of mental health outcomes among adolescents exposed to Adverse Childhood Experiences (ACEs). It would also allow for verification of the temporal sequence among ACEs, moderating factors, and subsequent mental health outcomes.

2. Complex Model Building

Further, examine the potential interaction between family functioning and school mental health services. Although the current study did not identify significant moderating effects for these factors individually, exploring their combined influence may reveal synergistic protective mechanisms against the detrimental effects of ACEs. Such insights could contribute to a more comprehensive theoretical foundation for designing targeted interventions.

3. School-Family Partnership Intervention and Evaluation

Develop and implement structured school–family partnership programs—such as collaborative mental health education and coordinated parenting workshops—and conduct rigorous follow-up evaluations. These efforts would help assess the real-world effectiveness of systematic school–family collaboration in improving mental health outcomes among adolescents with ACEs, thereby bridging the gap between theoretical recommendations and practical implementation.

Works Cited

Balistreri, K. S., & Alvira-Hammond, M. (2016). Adverse childhood experiences, family functioning and adolescent health, and emotional well-being**.** *Public Health, 132*, 72–78. https://doi.org/10.1016/j.puhe.2015.10.034

Kovács-Tóth, B., & Kuritárné Szabó, I. (2023). Az ártalmas gyermekkori élmények hatása a mentális és szomatikus egészségre gyermek- és serdülőkorban. *Orvosi Hetilap, 164*(37), 1447–1455. https://doi.org/10.1556/650.2023.32851

Bomysoad, R. N., & Francis, L. A. (2020). Adverse childhood experiences and mental health conditions among adolescents. *Journal of Adolescent Health, 67(6),* 868–870. https://doi.org/10.1016/j.jadohealth.2020.04.013

Craig, F., Servidio, R., Calomino, M. L., Candreva, F., Nardi, L., Palermo, A., Polito, A., Spina, M. F., Tenuta, F., & Costabile, A. (2023). Adverse childhood experiences and mental health among students seeking psychological counseling services. *International Journal of Environmental Research and Public Health, 20*(10), Article 5906. https://doi.org/10.3390/ijerph20105906

O'Reilly, M., Svirydzenka, N., Adams, S., & Dogra, N. (2018). Review of mental health promotion interventions in schools. *Social Psychiatry and Psychiatric Epidemiology, 53*(7), 647–662. https://doi.org/10.1007/s00127-018-1530-1

Shu, Z., Chen, S., Chen, H., Chen, X., Tang, H., Zhou, J., Tian, Y., Wang, X., & Zhou, J. (2024). The effects of subjective family status and subjective school status on depression and suicidal ideation among adolescents: The role of anxiety and psychological resilience. *PeerJ, 12*, e18225. https://doi.org/10.7717/peerj.18225

Parrigon, K. L. S., & Kerns, K. A. (2015). Family processes in child anxiety: The long-term impact of fathers and mothers. *Journal of Abnormal Child Psychology, 44*(7), 1253–1266. https://doi.org/10.1007/s10802-015-0118-4

Wang, D., Jiang, Q., Yang, Z., & Choi, J.-K. (2021). The longitudinal influences of adverse childhood experiences and positive childhood experiences at family, school, and neighborhood on adolescent depression and anxiety. *Journal of Affective Disorders, 292*, 542–551. https://doi.org/10.1016/j.jad.2021.05.108

The Need for Increased Awareness and Institutional Support for Strabismus By Sungwoon Jung

Abstract

Strabismus is an eye condition present in approximately 2-4% of children ("What Is Strabismus?"). It is often misunderstood by many people as a mere aesthetic issue rather than a legitimate, serious physiological disability. While people simply try to minimize the condition and turn a blind eye to it, implications and ramifications quietly continue to bubble under the surface. Although it can be treated in the early stages of life with medical means such as eye patching and correction therapy, many children experience difficulty due to a lack of proper institutional support for this condition. This is significant because the problem not only shapes early life experience but can have lasting repercussions into adulthood as well. This paper explores some of the disadvantages faced by people born with strabismus, which include hardship within school and social life, perceived intellectual deficiency, and future impact on overall life satisfaction. By assessing the current status of strabismus and comparing it to a different disorder, I will underscore the need for increased awareness of the condition on a more institutional scale and further recognition of strabismus as a legitimate disability.

Introduction

Eyes are arguably one of the most important features of a human being. Indeed, they, and the skill they provide, are part of the core five senses we as humans rely on to move and function throughout the world. Daily activities such as walking, cooking, and speaking with others become extremely challenging without the support of sight. Though undeniably important, the gift of properly functioning eyes may even perhaps be taken for granted.

Along with practical utility, a significant amount of aesthetic emphasis is put on the eyes. Eye colors such as blue eyes, green eyes, and hazel eyes have colloquial cultural and social connotations in Western culture, often being described as desirable eye colors. Apart from the actual eyeball itself, the general eye shape is also seen to affect one's appearance to the point where surgical intervention for aesthetic purposes is widespread in countries such as South Korea. Lastly, the prevalence of phrases such as "eyes are the window to the soul" signifies that eyes, and the overall impression they give, are defining features of a human being that greatly affect how we view ourselves and how others view us.

Because they hold such an important role, damage to the eye can be very consequential from both a psychological and a physiological standpoint. Even something as minuscule as a speck of dust in your eye can make it immediately water and lead to a feeling of discomfort, so it is no wonder that various medical problems can arise in that area, which warrant specialized care. Because eyes are indispensable and slow to regenerate (if at all) once damaged, we approach them with a different level of delicacy and caution as opposed to, for example, bones that can break but heal with time as long as they are secured in place. Within the multitude of eye-related issues that exist, the main focus of this paper will be specifically on strabismus and its ramifications.

Strabismus is a disorder in which both eyes do not line up in the same direction, as opposed to the proper functioning (orthotropic) state, where both eyes are trained on the same object or location. Therefore, they do not look at the same object at the same time. One of the most common and widely recognized forms of strabismus is colloquially known as "crossed eyes," where both eyes turn inward ("What Is Strabismus?"). Normally, the two eyes work together to form a coherent image that gets sent to the brain. However, with strabismus, there is an imbalance that leads to misalignment. Other forms of strabismus include each eye turning in distinct directions, only one eye experiencing strabismus, or intermittent bouts of strabismus. Effects of strabismus can range from being minimal (as the eye adjusts to the imbalance and one starts to overcompensate for the other), to blurred vision, headaches, and more ("What Is Strabismus?"). These conditions can be cured in the early stages of life with non-surgical practices such as glasses or eye patches if addressed early. Furthermore, surgical intervention is also possible, with an over 70% overall success rate in children aged 1-6 years old (Hinterhuber). Although a condition affecting 2-4% of the population may seem like an insignificant percentage, in the United States, that translates to millions of children each year.

Peer Exclusion due to Strabismus

Strabismus produces difficulty for the individual experiencing the vision problems themselves, but can also be a source of friction for the external observer as well. Although people like to think that they treat everyone equally and do not form premature judgements, in reality, that is commonly not the case. This is where another significant issue arises. Children as well as adults with strabismus can be disproportionately prone to mental disorders and difficulties due to negative perception from others. Particularly for developing elementary and middle school students who have not fully grasped the potential consequences of how certain actions may affect other people, any thoughts of "otherness" may show through unfiltered—even if they do not intend to do so maliciously. On the other hand, childhood is a highly sensitive time for children and adolescents. During this time, students aim to attain a strong sense of belonging, and thus a social goal may be prioritized over an academic one ("Characteristics Middle Grade Students"), and so the likelihood of intentional malice is definitely still possible. From this lens, there is a strong potential for hostile physical or psychological bullying to occur. Indeed, children who were perceived to be less physically attractive, overweight, or have a sight/hearing/speech disability were more likely to experience bullying and teasing from peers (Sweeting and West). Since children with strabismus are perceived to have distinct visual differences, they are indubitably included in this group as well. Past research shows that children with strabismus were rated more negatively compared to orthotropic (normally aligned) peers on measures including intelligence, health, sociability, and satisfaction in life (Mojon-Azzi et al.). Moreover, it was also found that children with strabismus started experiencing these social ramifications, such as being less frequently invited to birthday parties from around the age of six to right around the age where peers start noticing differences in themselves and those with strabismus (Mojon-Azzi et al.). These observations indicate that prejudice and discrimination from

abnormal eye alignment begin at a young age and thus have even longer to become fully embedded into a child's psyche. Regardless of whether the feeling of "otherness" is intentionally antagonistic or not, the receiving end will nevertheless feel the effects.

Long-Term Psychological and Developmental Impact on Children with Strabismus

Time in school is a critical developmental period that may strongly influence the paths of our society's future adults. Because it is such an important period, social exclusion and disadvantages can result in long-term psychological hurdles, even when children fully develop and become adults. Studies have found that early bullying experiences are connected to long-term psychological problems, such as depression, anxiety, phobias, and risks of suicide (Arseneault). Children with strabismus are also victims of this phenomenon. Researchers found that 41.3% of children with strabismus are diagnosed with mental illnesses in the early stages of their lives, which is considerably more than that of their peers; those with intermittent exotropia, a condition of strabismus where both eyes turn outward intermittently, are at a bigger risk, showing higher rates of psychiatric problems, suicidal thoughts, and visits for mental health therapies (Mohney, et al.). Moreover, children with strabismus are found to experience higher levels of social problems, including separation anxiety, social dissatisfaction, and low self-esteem (Cumurcu, et al.). These early-stage issues remain, or even worsen, outside of the classroom as they become obstacles for peer relationships, academic achievements, and confidence in social participation. In the same way victims of weight-related bullying experience poor romantic relationships and life satisfaction, victims of eye-alignment bullying may face those kinds of problems even after finishing school and carry the additional stresses and wounds with them into their futures as they become functioning and contributing adults to society (Stamate et al.). But the difference between people with strabismus compared with those with other publicly recognized conditions is that they have to endure this social exclusion while receiving far less support and protection from institutions.

Overlook of Strabismus in U.S. Education

Despite affecting a non-negligible number of children, strabismus is largely dismissed within the context of inclusive education in the United States. Although efforts to support children with disabilities, such as federal education laws like the Disabilities Education Act (IDEA), exist, they often mention visual impairment in relation to more extreme conditions, such as total blindness, rather than more nuanced disabilities such as strabismus (20 U.S.C 1401). Thus, strabismus is put on the back burner, which leads to a lack of proper planning to support children with this condition. When the existence of strabismus is not familiarized to youths and adults as a very real possibility, it is not surprising that feelings of discomfort may arise when they are eventually encountered. It is natural for stronger public advocacy and social recognition to be focused towards readily recognizable and severe disabilities, but conditions such as strabismus, which can seem comparably less urgent, must not be neglected. While it is a great thing that some disability awareness has become more emphasized than in the past, that also

simultaneously can mean other existing disabilities become relatively deemphasized and overlooked. As a result, children with strabismus are often glossed over and rarely provided with Individualized Education Programs (IEPs) or even proper recognition that strabismus can significantly affect their development during school life (Williams et al). An IEP is a legal document in the United States provisioned under the Individuals with Disabilities Education Act (IDEA) to support public school children who have the right to receive special education. This document guides important figures such as schoolteachers and parents in providing adequate education and support for the child from a professional standpoint, so it is an essential blueprint in ensuring the proper, structured education and care of a child with a disability. Teachers need the guidance of this document as the teacher-student relationship is crucial for a student's life at school. A study on the student-teacher relationship finds that a more supportive approach from teachers is linked to higher achievement and engagement of students in school; thus, proper care and attention from teachers can make a big difference in a child's life (Cornelius White). As the World Health Organization guidelines point out, community-based rehabilitation is important for the successful development of children with disabilities, and these include supporting key stakeholders, which include policy makers as well as school leadership, in meeting the basic needs of the disabled and encouraging the empowerment of the disabled and their families (World Health Organization). The lack of social recognition of strabismus is a core problem that enables the continued prolonging of its minimal support, leaving children feeling invalidated during the growing stages of their lives.

Lack of Institutional Recognition of Strabismus as a Disability

It is known that strabismus can be treated during early stages of developmental periods, but the lack of its recognition as a disability from the government makes it difficult for children and their families to get help in any form. Strabismus is not always recognized or accepted when applying for disability-related aids, even though it can critically damage a person's quality of life. Financial aid for medical treatments such as corrective vision surgeries is often unsupported unless it is a surgery that restores vision for both eyes, which indicates that strabismus is not considered a condition that warrants treatment unless it passes a certain socially determined threshold of impairment (Chawla, et al). This rejection essentially misses the golden opportunity to cure strabismus in its early, initial stages for children, where costs and treatment plans may be considerably more manageable from a medical, logistical, and financial perspective. As this condition is left untreated, severity may increase along with additional risks of lasting psychological and social concerns. In contrast, other major publicly recognized conditions, such as autism, speech impairments, or hearing disabilities often receive much better treatment support from the government (Taylor, et al). Government programs like Medicaid and IDEA provide financial aid for programs such as speech therapy for speech impediments, psychological treatments for autism, and hearing aids for hearing disabilities, service options which are not present in the same quality for those with strabismus (Marsh). Furthermore, many private medical insurance companies now cover treatment for autism, a fact that was not true in the past

but, through continued efforts for increased awareness toward the issue, was eventually attained (Candon, et al.). The success case of increased recognition of autism throughout the years suggests the promising possibility that strabismus' public perception can be changed as well. Overall, the lack of inclusion of strabismus in disability policy contributes to low levels of support not only on an individual level, but on a more systematic degree as well.

Autism as a Success Case Model

Autism spectrum disorder (ASD) shows us an example of how improved public perception and recognition can lead to better systematic support. From when autism was first introduced as a concept in the early 1900s to how it is dealt with currently, understanding of the disorder has gone through multiple radical changes; it started off as a term used similarly to schizophrenia, and now it is known as a neuro-cognitive/developmental disorder (Evans). Similar to current problems with strabismus, there was a lack of public knowledge about autism, and it was not included in school education and institutional support. It was not until 1959 that the Mental Health Act was passed in Great Britain, which served as a catalyst for the change in perceptions regarding the condition (Evans). Bringing the condition into increased publicity and support from an institutional level began to make shifts in a positive direction. Continued public efforts, such as publicity campaigns like the establishment of World Autism Awareness Day in 2007, helped bring many people's attention to this issue globally, with these efforts still continuing to this day. Indeed, recent studies conducted in Brazil have shown that autism awareness campaigns can significantly evoke public engagement, which makes information about autism more easily accessible (Costa et al.). These campaigns have not only significantly helped to reduce negative public perception of autism but also improved explicit public support for this issue. The relationship between public knowledge of a disability and corresponding support is reflected in research; higher public awareness was correlated with more inclusive attitudes and approaches towards individuals with disabilities (Wang, Ziru, et al). The Combating Autism Act, passed in 2006 and subsequently amended to the Autism CARES Act in 2014, displayed continued care and improvements in ongoing support for autism and its treatment in an extensive way by strengthening federal connections, which shows a direct example of a link between increased advocacy leading to an improved institutional effect. Although increasing cultural awareness may seem like a small first step, it is essential to demonstrate that people truly care about it. Only then can it attract the pragmatic support, such as funding and research, necessary for meaningful and long-lasting progress. This is the domino effect we aim to initiate. While autism spectrum disorder and strabismus do have notable differences, these efforts and past successes nevertheless give us a coherent, concrete example for strabismus to hopefully follow in its path: increased public awareness can lead to significant improvement in treatments and support for people with the condition.

Conclusion

Although strabismus affects millions of children in the United States, it remains overlooked in both personal and collective aspects compared to some other disabilities. Lack of proper recognition and help from society leaves many children with strabismus neglected, which not only allows room for the eye misalignment issue to worsen but also leaves the psychological challenges they face from social exclusion and prejudice untreated. The studies reviewed in this research paper show that strabismus in childhood and related social rejections can lead to serious future disadvantages, including peer bullying and perceived intellectual inferiority, to long-term psychological problems. It is important to emphasize that the condition is highly treatable, especially in the early stages of life, and, with proper support, many of its long-term effects can be significantly reduced or avoided. However, if left unaddressed, it may lead to serious and lasting impacts. Therefore, it is crucial that it not be overlooked, as early intervention can make a meaningful difference.

Past successes of other conditions in bringing public awareness and institutional recognition imply that the challenge of strabismus can also be met with similar efforts. They include raising awareness of strabismus as a legitimate but treatable disability, increased exposure and training regarding strabismus in school environments, and more government-backed access and support for early treatments like eye patches and corrective surgery. While it is possible to focus on the medical solutions, they can be largely inaccessible to many due to financial barriers. The effects of implementing the solutions through public channels, such as awareness programs and social policy awareness on a larger scale, will be much more equitable and far-reaching than simply reducing financial barriers. With the coordination of help from society, tangible and practical progress can be made in addressing the current problems stemming from the eye condition. Shifting social perception of strabismus as a mere cosmetic issue to recognizing it as a serious, legitimate disability is a key starting point to solving these problems, as they can act as the catalyst for the domino effect. Ultimately, with these efforts, we can get one step closer to the goal of children with strabismus receiving early treatments, avoiding and minimizing unnecessary psychological harm, and finally, growing into healthy adult members of society.

WORKS CITED

Arseneault, Louise. "The Long-Term Impact of Being Bullied in Childhood." *Current Opinion in Psychiatry*, vol. 30, no. 6, Nov. 2017, pp. 440–445. doi:10.1097/YCO.0000000000000355.

Candon, Matthew K., et al. "Insurance Mandates for Children's Autism Increase Service Use, but Do They Also Increase Out-of-Pocket Spending?" *Pediatrics*, vol. 143, no. 4, Apr. 2019, e20182040. *PubMed Central*, https://doi.org/10.1542/peds.2018-2040.

"Characteristics of Middle Grade Students." California Department of Education, 2012.

Chawla, O., et al. "Understanding Parental Hurdles in Accessing Strabismus Treatment in Children." *Ophthalmic Epidemiology*, vol. 31, no. 5, 2024, pp. 429–437. *ScienceDirect*,

Cumurcu, T., B. Elbozan Cumurcu, Ö. Ozcan, S. Demirel, C. Duz, E. Porgali, and S. Doganay. "Social Phobia and Other Psychiatric Problems in Children with Strabismus." *Canadian Journal of Ophthalmology*, vol. 46, no. 3, June 2011, pp. 267–270. doi:10.1016/j.jcjo.2011.05.017.

Cornelius White, Jeffrey. "LearnerCentered TeacherStudent Relationships Are Effective: A MetaAnalysis." *Review of Educational Research*, vol. 77, no. 1, 2007, pp. 113–143. *JSTOR*, www.jstor.org/stable/4624889, https://doi.org/10.2307/4624889.

Evans, Bonnie. "How autism became autism." *History of the Human Sciences*, vol. 26, no. 3, 8 May 2013, pp. 3–31, https://doi.org/10.1177/0952695113484320.

Hinterhuber, Laetitia, et al. "Postoperative outcome and influencing factors of strabismus surgery in infants aged 1–6 years." *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 262, no. 7, 16 Feb. 2024, pp. 2299–2307, https://doi.org/10.1007/s00417-024-06404-1.

Marsh, Ian B. "Adult Strabismus Surgery Is Not Merely Cosmetic." *Eye*, vol. 28, 2014, pp. 1036–1037. Springer Nature, https://doi.org/10.1038/eye.2014.136.

Mohney, Brian G., Jeff A. McKenzie, Joseph A. Capo, et al. "Mental Illness in Young Adults Who Had Strabismus as Children." *Pediatrics*, vol. 122, no. 5, Nov. 2008, pp. 1033–1038. doi:10.1542/peds.2007-3484.

Mojon-Azzi, Stefania Margherita, et al. "Strabismus and discrimination in children: Are children with strabismus invited to fewer birthday parties?" *British Journal of Ophthalmology*, vol. 95, no. 4, 18 Aug. 2010, pp. 473–476, https://doi.org/10.1136/bjo.2010.185793.

Mojon-Azzi, Stefania Margherita, et al. "The perception of strabismus by children and adults." *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 249, no. 5, 10 Nov. 2010, pp. 753–757, https://doi.org/10.1007/s00417-010-1555-y.

Stamate, Iris Fibia, et al. "Association between Weight-Related Bullying and the Academic, Professional, and Romantic Pathways in Emerging Adulthood." *Research Square (Research Square)*, Research Square (United States), Nov. 2024, https://doi.org/10.21203/rs.3.rs-5348071/v1. Accessed 19 Aug. 2025.

Sweeting, Helen, and Patrick West. “Being different: Correlates of the experience of teasing and bullying at age 11.” *Research Papers in Education*, vol. 16, no. 3, Jan. 2001, pp. 225–246, https://doi.org/10.1080/02671520110058679.

Szlamka, Zsofia, et al. “The Role of Advocacy and Empowerment in Shaping Service Development for Families Raising Children with Developmental Disabilities.” *Health Expectations*, vol. 25, no. 5, Oct. 2022, pp. 1882–1891. *PubMed Central*, https://www.researchgate.net/publication/360945550_The_role_of_advocacy_and_empowerment_in_shaping_service_development_for_families_raising_children_with_developmental_disabilities

"S.843 - 109th Congress (2005-2006): Combating Autism Act of 2006." *Congress.gov*, Library of Congress, 19 December 2006, https://www.congress.gov/bill/109th-congress/senate-bill/843.

Taylor, Julie Lounds, et al. “Improving Parents’ Ability to Advocate for Services for Youth with Autism.” *Journal of Autism and Developmental Disorders*, vol. 53, no. 11, Nov. 2023, pp. 4307–4318. *PubMed Central*,

United States, Congress, *Individuals with Disabilities Education Act*. 1975.

Wang, Ziru, et al. “Factors Associated with Public Attitudes towards Persons with Disabilities: A Systematic Review.” *BMC Public Health*, vol. 21, no. 1, 3 June 2021, bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-11139-3, https://doi.org/10.1186/s12889-021-11139-3.

“What Is Strabismus?” *Cleveland Clinic*, 20 June 2023, my.clevelandclinic.org/health/diseases/strabismus-eye-misalignment.

Williams, Marian E., et al. *Quality of Individualized Education Programs for Children with Visual Impairments and Autism Spectrum Disorder*. Journal of Visual Impairment & Blindness, vol. 117, no. 3, May–June 2023, pp. 199–211.

World health organization. “Community-Based Rehabilitation: CBR Guidelines.” *Www.who.int*, 12 May 2010, www.who.int/publications/i/item/9789241548052.

Advances and Challenges in Kidney Regeneration: Focusing on Vascularization in Tissue Engineering and Regenerative Medicine By Aungsula Pathak

Authors and Affiliations:
Author: Aungsula Pathak
I would like to acknowledge Dylan Yeo (University of Oxford) for his mentorship during my writing.

Abstract

Chronic kidney disease (CKD) often progresses without noticeable symptoms and becomes clinically apparent only at advanced stages, when renal function is already severely impaired. By the time of diagnosis, kidney structure is frequently irreversibly damaged, leaving many patients dependent on kidney transplantation. However, transplantation is limited by the scarcity of donor organs and long waiting times. Although advances in immunosuppressive therapies have improved short-term outcomes, long-term success remains affected by chronic rejection and the ongoing burden of immune management. In response to these challenges, regenerative medicine has emerged as a promising alternative approach. Kidney tissue engineering, in particular, aims to create functional renal tissue using strategies such as biological scaffolding, bioprinting, and microfluidic "chip-based" models.

A major challenge in renal tissue engineering is achieving effective vascularization. The kidney's highly complex vascular system is essential for filtration, oxygen delivery, and overall renal function, yet replicating this network in engineered tissues remains difficult. Current research efforts to restore renal vascular structures include approaches such as decellularized organ scaffolds, three-dimensional (3D) bioprinting, and microfluidic models. While decellularized scaffolds show potential for preserving native vascular architecture, they continue to face significant limitations, including thrombosis and insufficient blood flow following transplantation. Additionally, many bioprinting techniques that incorporate angiogenic factors or directly print vascular networks are still in early stages of development and require further optimization.

This review examines current strategies in kidney tissue engineering with a specific focus on vascularization, highlighting recent advances, ongoing challenges, and future directions in the field.

Methods- Review; Sources were gathered from google scholar and PubMed.

Introduction

Background and Context

Chronic kidney disease (CKD) is a major cause of death that currently affects 10-13% of the global population and experiences an upward trend in prevalence[1]. CKD can be caused by a

number of underlying conditions, including a history of diabetes, hypertension, autoimmune diseases like lupus, HIV infections, primary glomerulonephritis, and more[2]. It is most common among adults aged >60 years and has recently shown a rising trend in the United States population[3]. As such, CKD has shown disconcerting correlation with obesity and substance misuse—likely factors contributing to the upward projection.

CKD impairs and gradually deteriorates essential processes regulated by the kidneys such as waste removal, body-fluid balance, blood pressure regulation, and vitamin metabolism[4]. What's particularly dangerous in these cases is that these symptoms are hard to detect early on. Meaning, the effects must reach a severe point to be detected. With that being said, while not deadly alone, these CKD-related complications can worsen over time to much more dangerous outcomes, including heart disease, anemia, metabolic acidosis, bone disorders, among others—primarily characterized by hormone dysfunction and waste accumulation. This places patients at risk for more serious conditions like multiple organ dysfunction syndrome (MODS)—where two or more independent organs fail to function properly—and significantly lowers their chance of recovery[5]. Most problematic of all, however, is the timeline in which CKD presents. CKD is defined by the presence of kidney damage or an estimated glomerular filtration rate (eGFR) below 60 mL/min/1.73 m² that lasts for 3 months or longer, regardless of the preliminary cause[6]. Symptoms indicating that medical attention may be required include one or a combination of: frequent, foamy urination; swelling of the limbs; itchy, dry skin; bad breath—possibly of a "fishy" or metallic taste—; and sudden weight loss[4]. However, as aforementioned, these symptoms are difficult to detect early on.

Primarily asymptomatic in its early stages, CKD becomes apparent only after slow evolution, at which point the disease is oftentimes irreversible. It's crucial not to overlook urgent care, even after months of neglect, as CKD will continue to progress regardless of lifestyle improvements. Although studies examining healthy dietary patterns, such as the "Mediterranean Diet," have demonstrated a reduction in the risk and progression of CKD, certain non-modifiable factors, such as being male, older age, and non-White ethnicity, will continue to negatively progress CKD anyway[7, 8, 9]. This means that certain demographics—people of color, males, and people aged over 60—are at greater risk for CKD and CKD progression. But what does that mean for the kidney?

With continued insult, there is loss of renal architecture and functional recovery is no longer possible. By this time, most patients have no choice but to undergo transplantation as the sole means of recovery. Success of kidney transplantation, however, is constrained by availability of donor organs and waiting time. In fact, according to a meeting report provided by Harvard Medical School in 2003, from the median waiting time for kidney transplantation that took 19 months a decade ago, patients on the list today have to wait over 3 years due to the growth in the waitlist from about 15,000 patients in 1990 to 55,000 patients by mid-2002[10]. Despite this, the number of annual cadaveric kidneys transplanted has remained the same at about 9,000 annually.

This clearly indicates that kidney transplantation is such a timely process to which there is a high chance for deterioration among CKD patients. And, as mentioned earlier, the amount of people suffering from CKD is continuously increasing—only adding to the discrepancy. Also, renal transplantation and immunity management after the surgery is faulty even in case of successful transplantation.

Improvements in immunosuppressive therapy have lowered the rate of acute rejection to some extent and improved moderate-term outcomes; however, the effectiveness of these therapies has been limited due to unclear relationships between the concentration of the drugs and their outcomes, and also by chronic transplant rejection which may be presented long after initial success[11]. These issues significantly affect morbidity in individual transplant recipients and thus should be more effectively addressed. The recognition of the individual nature of the immune system in each transplant recipient has elicited interest not only in personalized immunosuppressive treatments but also in another area of medicine: regenerative medicine.

Generally, regenerative medicine approaches aim to take cells from the patient and develop a completely new, but individual, organ that will replace the previously damaged one, with reconstitution of function, and without immune rejection and also ensuring longevity. Stem cells play a critical element in regenerative medicine where the cells act as a source of progenitor or undifferentiated cells for cell replacement therapies[12]. Some of the significant themes up to now in regenerative medicine, through ex vivo and in vivo techniques alike, involved transplantations of stem cells, 3-D bioprinted organoids, and treatments of antibodies[13].

One part of regenerative medicine, which has lately gained attention, includes tissue engineering as a new and effective cluster of methods with potential in the area of personalized medicine and reconstruction. More specifically, tissue engineering evolved from biomaterials development and is defined as the combination of scaffolds, cells, and biologically active molecules in an attempt to create functional tissues[14]. Tissue engineering seeks to fabricate functional constructs to restore, maintain, or improve diseased tissues or entire organs—an ideal goal for CKD patients when considered as an alternative to transplant waiting lists and organ donors.

However, due to the size and complexity of the kidneys, like in many other bioengineering processes, a major challenge in developing clinically relevant tissue-engineered kidneys is achieving precise vascularization—which is integral to replicating proper kidney function—posing the question: **What are some common techniques for kidney regeneration, and subsequently, how effective are they in solving challenges related to vascularization?**

For anatomic background, the kidneys are described as bean-shaped organs that weigh about 150 to 200 g in males and about 120 to 135 g in females[15]. Each kidney measures about 10 to 12 cm in length, 5 to 7 cm in width, and 3 to 5 cm in thickness—approximately the size of a closed fist. They are located retroperitoneally on either side of the posterior abdominal wall, with the right kidney being slightly lower than the left because it lies by the liver[16.] Now, with their size and

critical location next to several organs, including the liver, colon, pancreas, spleen, and stomach, alone makes the kidneys very difficult to regenerate. With the addition of their complicated vasculature, it is only more so.

The most complex and important feature of the kidney is its vasculature, which is considered the most difficult aspect in engineering kidney tissue. Every kidney has a large renal artery and renal vein that split into smaller vessels. These incorporate interlobar arteries that run along from the renal artery. These interlobar arteries then divide into arcuate arteries, which make up a system that allows for the blood flow through the extremely stratified organ 17]. The anatomy of this vascular system is even further complicated due to the different anatomies that appear during the kidney's embryogenesis and are characterized among different subjects. This also makes it hard for both kidney transplantation and tissue engineering because the vascular network of each kidney has to be critically assessed and matched to avoid fatal outcomes in surgical procedures[18]. In the case of transplantation, matching can indeed take months due to the long-term effort of finding a donor with compatible vasculature, further emphasizing the role vascular assessment plays in transplant success. Vascularization in the kidney is important not only structurally, but also physiologically, especially in blood pressure regulatory mechanisms and fluid exchange for waste excretion. Indeed, the relationship is demonstrated by the juxtaglomerular apparatus (JGA)—extremely specialized structure with proven significance in the regulation of renal blood flow and filtration rates[19].

The kidney has cells in the afferent arterioles of the JGA that, if blood pressure is low, will secrete a hormone called renin. This will cause the arterioles to dilate, thus increasing blood flow. The other component of the JGA is the macula densa that monitors sodium concentration in the fluid passing through the kidney. If sodium levels are low, the JGA signals the arterioles to widen, increasing blood flow and filtration. If sodium levels are high, the arterioles tighten, reducing blood flow and filtration. This regulation ensures that blood flow and filtration within the kidney are optimized to maintain homeostasis, allowing the kidneys to efficiently filter waste, excess fluids, and toxins from the blood. Therefore, without proper vascularization, the filtration capacity of the kidney would be compromised, and then the organ would serve no purpose at all. Again, the main functions of the kidney are removal of waste, maintenance of body-fluid balance, regulation of blood pressure, and vitamin metabolism; if the vasculature is not correct, none of these important functions can take place, showing how important the relationship between kidney vasculature, function, and overall homeostasis is.

In printing a kidney, there are certain challenges that must be overcome to achieve a functional organ. Among the primary challenges is how to differentiate stem cells into the expected cell types of a kidney. This can be achieved prior to and after printing the cells. In some cases, the stem cells are pre-differentiated into some of the cell types of the kidney, such as epithelial cells, endothelial cells, or podocytes, before incorporation into the bio-ink. In another approach, the stem cells can be printed in an undifferentiated state and then stimulated to differentiate within

the printed scaffold through the addition of selective growth factors, signaling molecules, or by creating suitable biochemical cues in the environment. Timing and technique of differentiation depend on the outcome required and the availability of the bio-printing system.

The second major challenge lies in the manner in which scaffolds are employed to support the cells and enhance formation of a "classic kidney" form. Scaffolds are biomaterials, usually biodegradable polymers or natural matrices like collagen, that provide a temporary shape to which cells attach, grow, and organize to generate the required tissue structure. These scaffolds guide the cells into the correct form and provide the mechanical support that the tissue requires to grow. The scaffold will degrade, and the cells will assume the structural roles and naturally grow the tissue. The design of scaffolding is important as it must mimic the extracellular matrix (ECM) of the kidney to allow for correct cell alignment and organization.

Finally, making the cells grow at the right place and to actually perform what they would naturally is another challenge for kidney bioprinting. This is achieved by a combination of precise 3D printing techniques and using some specific bio-inks that allow different types of cells to be placed in their right anatomical positions. In addition, the in- and post-printing incubation conditions of the cells also have a very important role to play in guiding cell function. This includes the provision of suitable biochemical signals, oxygenation, and nutrition for the sustenance of cell differentiation, proliferation, and function. By controlling the variables of the cell types with high precision, researchers can cause each cell type to not only develop in the correct position but also to express its normal kidney-specific functions, i.e., filtration, reabsorption, and secretion.

Significance & Objective

Efforts to regenerate and implant simpler hollow structures like bladders, blood vessels, urethras, and parts of the upper airways have been successful in the short to mid-term[20]. However, bioengineering complex solid organs with complex vasculature, such as the kidney, has greater challenges and requires many different strategies to achieve success. No study has yet to accomplish this; however, **this review explores the different methods attempted thus far and evaluates their success in replicating kidney vascularization due to its importance in kidney function.**

Literature Review

For simplicity, this review will combine and review the most common bioengineering techniques that have been researched for kidneys thus far: (1) Biological Scaffolding Using Decellularization Techniques, (2) Bio-Printing, and (3) Kidney-on-a-Chip models. It is worth noting that these techniques are not used as a single method; they usually involve additional methods such as mechanical stimulation, cell deposition, and cell culture. These supplementary techniques are, however, not discussed in detail in this review, as the main focus is on evaluating

the effectiveness of the primary technique and its ability to achieve vascularization.

1. Biological Scaffolding Using Decellularization Techniques

The process of biological scaffolding through decellularization is where the cellular material of a donor organ is removed while the extracellular matrix (ECM)[21]. The premise is to uphold the structure of the ECM where new cells will be deposited. Here, the ECM serves as a scaffold for recellularization in which different cell types, including renal or endothelial cells, are seeded onto the scaffold to support the regeneration of viable tissue. Decellularization of kidney regeneration begins by removing a kidney, which may either be from an animal or a cadaver, before it undergoes detergent and enzymatic buffer processing to wash out cellular material. The resulting scaffold is then seeded with kidney-specific cells, stem cells, and endothelial cells to create functional renal tissue and vasculature. The recellularized scaffold is cultured under controlled conditions to enable cellular proliferation and differentiation to increase the cells from the limited implanted to a whole organ. The process is highly advantageous because it preserves the native architecture of the organ, which provides vital signals for cell attachment and differentiation. Decellularization is also useful for developing vascular networks as endothelial cells are able to graft onto the scaffold and have the potential to re-form vascularization required for kidney function.

The process does present complications, however. One significant limitation involves the source material since the procurement of sufficient and suitable decellularized kidneys is problematic, especially in the issue of organ donors' availability. Besides, even though recellularization procedures can occur, they often do not achieve successful production of fully functional tissues. The development of a full and functional vascular system still remains a major challenge.

In a study by Sullivan et al. decellularized kidney scaffolds have been successfully prepared from porcine kidneys while preserving the preexisting organ architecture, in particular, the vascular network from both the arterial and venous sides[22]. In this study, transplantation of a decellularized kidney into a pig was performed without recellularization in order to determine its ability to integrate with the host vascular system and to reestablish blood flow. Adequate blood perfusion and stable blood pressure were obtained, even in the absence of cellular components in the matrix. In the acellular matrix, though, there was significant thrombosis within minutes, characterized by infiltration of inflammatory cells and formation of thrombi. Although blood perfusion and blood pressure were maintained, thrombosis kept appearing, thus indicating that vascularization was still quite limited in this regenerative strategy. Another study by Katari et al., used a "cell-on-scaffold seeding technology" based on the very widely-accepted knowledge that cells require a support structure to proliferate, stay alive, and function[23].

In another study done by authors Katari et al., researchers acknowledged the potential for using human kidneys for scaffolding in bioengineering new kidneys. In a prior study they conducted, they found that decellularized pig kidneys to provide renal extracellular matrix (ECM) scaffolds

that preserve their gross architecture and structural elements, support cell proliferation in vitro and in vivo, and preserve a functional vasculature that is able to handle physiological blood pressure. However, more promising results were found in human kidneys. They found that detergent-based decellularization of discarded human kidneys could maintain their original ECM organization, biochemical characteristics, and angiogenic potential, as well as a viable vascular network—which is integral to this study. Moreover, their method removed immunogenic antigens—foreign substances that the body recognizes and responds to by producing antibodies—an important consideration in achieving the best clinical results and minimizing the likelihood of graft rejection in the long term. They do contend, however, that their current investigations are still very early stages and clinical translation is not immediately available in the short term. However, with further research, these studies have the potential to bridge the gap between bioengineering and transplantation.

2. Bioprinting

3D bioprinting involves the precise layering of cells and biomaterials to generate complex vascular networks within engineered constructs (vascular scaffolds, as explained above). The bioprinting technology has been used to develop renal in vitro models, enabling the creation of 3D constructs that combine cells, biomaterials, and other biological factors in pre-planned strategies[24]. The composition of the bio-material ink also plays a critical role in the success of the process. Bio-inks typically consist of living cells, which may be stem cells or tissue-specific cells, e.g., renal cells for a kidney. These cells are embedded in hydrogels like alginate, collagen, or polyethylene glycol (PEG), which present a hydrated microenvironment similar to that of natural tissue. Apart from these hydrogels, the bio-ink may be mixed with biomaterials like gelatin, cellulose, or silk fibroin to add structural support and promote cell growth. Growth factors or signaling proteins can be added to induce cell differentiation, and crosslinkers to stabilize the bio-ink after printing so that the printed structure does not collapse. All these components work together to mimic the physical and biological environment of the printed tissue.

Sterility is also necessary in the process of bioprinting because it deals with living cells that can be destroyed by contamination. Bio-printing is typically carried out in cleanroom laboratories that are to exacting standards, for example ISO 14644-1, to avoid particulate and microbial contamination. The bioprinting materials used, including the bio-inks, nozzles, and substrates, are all sterilized, for instance using autoclaving, UV or filtration, prior to use. Apart from this, sterile materials, gloves, and sterile hoods (like laminar flow hoods) are employed in preparation and assembly to avert contamination. Antimicrobial substances could also be included in some bio-inks, but these have to be wisely chosen so that they are not toxic to the cells. Sterility has to be maintained all along the process of bioprinting, starting from preparation to post-printing care, so that the printed organ can ensure viability and function.

This is an additive technology of manufacture that lays biomaterial inks in successive layers to form complex tissue. For the kidney, in order to create relevant models, cells, biomaterials, and/or bioactive molecules have to be chosen precisely. The 3D model generated through bioprinting is allowed to mature in a bioreactor or directly implanted in vivo[25]. In the last few decades, various printing methods have been developed with two primary techniques used most—extrusion or microfluidics-based methods—for creating kidney model[24]. One category of bioprinting, the pressure-assisted systems-being an extrusion-based technique, is the most commonly used among many research groups performing bioprinting due to its commercial availability at relatively cheaper costs[26].

Such systems usually comprise one or more cartridges that enable delivery of various combinations of cells and biomaterials. Plastic or glass cartridges are filled with the chosen biomaterial inks, and by applying gas pressure, material is extruded as a filament through a needle or nozzle into a scaffold or microfluidic channel system, thus creating an architecture of the intended organ—the kidney. An alternative approach, in this case, microfluidic-based method, pneumatic microdrop-on-demand bioprinting process utilizes compressed air to create and dispense precise microdroplets of bioink laden with cell onto appropriate substrates or scaffolds to be used for kidney applications[27]. This approach allows for the accurate placement of discrete droplets, where the process involves the filling of a reservoir with bioinks, enabling pneumatic pressure to dispel droplets at predetermined intervals through a nozzle.

One of the most important benefits of bioprinting is that it is extremely precise, thus allowing us to customize tissue architecture with particular cell types and materials. It has the ability to create extremely fine and reproducible models, offering a versatile way of helping to regenerate kidney tissue. Nevertheless, bioprinting also has its limitations, specifically when it comes to printing blood vessels. Researchers have printed vascular structures, but these networks often do not develop into fully working systems that can supply blood to larger tissues. The complexity of kidney blood vessels, which include small capillaries and complex tubule structures, makes it hard to copy using current bioprinting technology. Additionally, some have shown promising results in creating models of kidney tissue, though it is difficult to develop the cells fully and to get them to function properly, especially in more complex organs like the kidney.

Through the optimization of parameters including droplet size and printing speed, the spatial arrangement of cells and biomaterials can be improved, thereby fostering enhanced tissue functionality and bolstering overall cellular health throughout the printing process. While both types of bioprinting exhibit potential in tissue regeneration, the challenge remains in achieving adequate vascularization. Discussion: In terms of vascularization, the only significant research on bioprinting kidney vasculature has been seen in organ-on-a-chip models. Thus, bioprinting, while effective in tissue engineering, has little to no promising research in the regeneration of kidney vascularization via techniques outside of “chip models”—the next focus.

3. Kidney-on-a-Chip Models

As an extension of microfluidic bioprinting, organ-on-a-chip models are new possible targets for kidney regeneration. Traditionally, drug testing was performed using 2D methods for their simplicity and scalability. However, research shows that cells behave differently in 2D compared to 3D models, particularly in protein expression, drug response, cell shape, and migration; 3D models provide better representations of the physiological environment due to stronger cell–cell and cell–matrix interactions[29]. Organ-on-a-chip models are an example of this.

Microfluidics deals with the handling of small fluid volumes for applications such as DNA analysis, lab-on-a-chip systems, micropropulsion, and microthermal technologies[30]. Of interest, organ-on-a-chip techniques allow the development of the functional units of single or multiple organs—in this instance, the kidney. In the field of nephrology, renal tubular cells are incorporated into microfluidic chips in the development of kidneys-on-a-chip.

For instance, kidney-on-a-chip technology could enable real-time assessment of patient-specific drug responses in hospitals. The kidney-on-a-chip technology would not be inserted into the patients; it would be an external tool used for testing and research purposes, enabling one to analyze patient cells or tissues in a controlled environment to make informed treatment decisions without invasive procedures—accomplishing one of the benefits of regenerative medicine.

Looking at the more comprehensive structure first, the organ-on-a-chip is a microfluidic device and an extension of bioprinting[31]. The device contains perfused chambers with living cells arranged to mimic the physiology of tissues and organs. Numerous organ-on-a-chip systems have been developed, including models for gut, lung, blood vessel, cancer, bone marrow, and kidney-on-a-chip[32, 33, 34, 35, 36, 37]. Organ-on-chip models aim to mimic the physiological functions of human organs in vitro, with the goal of improving drug testing, disease modeling, and personalized medicine. Organ-on-a-chip engineering includes designing and fabricating a microfluidic device that mimics organ functions. Cells are harvested and added to the chip compartments, where a perfusion system mimics fluid dynamics to provide nutrients and remove metabolic waste products[30]. The device is often functionalized with extracellular matrix proteins to enhance cell adhesion, and cellular behavior is monitored in real-time to enable drug testing and disease modeling. In vivo, renal cells are exposed to complex and dynamic biomechanical signals that influence their behavior and function.

Mechanical stimulation has also been integrated into the kidney-on-a-chip models to recapitulate the biophysical microenvironment and study the response of the cells to various mechanical forces[38]. The kidney cells sense both the passive and active biomechanical stimuli. Passive forces originate from interactions with the cell-matrix interface arising from changes in the stiffness of the ECM, geometrical confinement, surface topography, or static stress. Active mechanical stimuli include fluid shear stress (FSS), tensile stretching, and compression that have a large range of frequencies, magnitudes, and durations, depending on the specific type of kidney cell[39].

Narrowing down to the kidney specifically, kidney-on-a-chip designs feature two compartments: a top channel that simulates the urinary lumen with fluid flow, and a bottom chamber that represents the interstitial space filled with media[40]. The goal of kidney-on-a-chip treatments is to create a realistic model for testing drug effects and studying kidney diseases. Authors Jang and Suh used rat distal tubular cells with shear stress around 1 dyn/cm², while Ashammakhi et al. utilized human proximal tubular cells to study cisplatin nephrotoxicity, with shear stress at approximately 0.2 dyn/cm²[41, 42]. These studies both demonstrate the potential of kidney-on-a-chip models in drug advancement and stress resistance but are both conducted on a smaller scale and do not attempt to revive larger portions of the organ.

Given the variety of cell types and functional units in the human kidney, a biomimetic kidney-on-a-chip should replicate key interactions, especially between glomerular endothelial cells and podocytes[43]. In addition, pressure differentials, the pattern of renal tubular segments, fluid flow dynamics, and cellular functionalities have to be considered. While different models of glomerular, proximal tubular, and distal tubular physiology have been developed, there has been no fully integrated kidney-on-a-chip system realized[21]. The kidney-on-a-chip technology can enable personalized drug testing and toxicity assessments in the near future, improving understanding of renal diseases and helping to develop treatment options tailored to target specific diseases, and it will open improved outcomes for patients with kidney diseases. As far as vascularization, an in vitro kidney depends on recapitulation of the renal endothelial–epithelial exchange interface—a great challenge that requires material selection and dimensioning with extreme care[43].

Recently, Rayner et al. proposed a new device called the human renal vascular–tubular unit (hRTVU), where a thin collagen membrane formed the interface. The method of using traditional membranes like PDMS and polycarbonate, which are unsatisfying to support effective cell integration, was replaced[44]. The device imitated the dynamic environments of the kidney by using a cell-remodelable hydrogel and controllable perfusion flow, thus allowing for an accurate quantification of albumin reabsorption. However, it suffered from issues due to a degrading flow profile and construction inaccuracy that prevented the attainment of vascularization at all. In contrast, Homan et al. created a more straightforward 3D-printed PDMS device using organoid aggregates in order to mimic functional complexity[45]. They were able to achieve perfusable vasculature with the help of the vasculogenesis model. They found that the flow induced vascularization as well as maturation of kidney organoids into tubular and glomerular regions. Despite this, the native vasculature was unable to support flow in the microvascular networks effectively, indicating the need for further improvements.

In this case, vascularization was not achieved either. As this was the last focus of this review, it is evident that there is not yet an alternative to organ donor transplants. Currently, no regenerative technique has fully mimicked kidney vascularization, particularly the complexity of the kidney's vasculature, which poses one of the greatest challenges in tissue engineering. While significant

progress has been made in creating vascular networks within engineered tissues, replicating the intricate and functional vascular structures required for kidney regeneration remains an unsolved problem. Much more research is required.

Discussion

Current Techniques in Vascularization

Several promising techniques try to solve the vascularization of kidneys but are ineffective yet. Scaffolding & Bioprinting: 3D bioprinting is an emerging technique that holds great potential in the creation of vascular structures. Scientists have been able to bioprint blood vessel-like structures using various types of bioinks and growth factors. Many studies have used bioprinting to fabricate vascular networks within skin, cartilage, and bone tissues. However, the biggest challenge remains—to create vascular structures that will functionally connect with the body's circulatory system for the fabrication of complex organs like the kidney. The printed structures to date all tend to lack full maturity or the ability to sustain perfusion in the long term without issues of thrombosis or poor perfusion.

Decellularization: This process removes cells from a donor organ, leaving an ECM scaffold that still maintains the architecture of the organ, including vascular structures. The idea is to repopulate this scaffold with the patient's cells (recellularization), which may reduce immune rejection. Decellularized kidneys have now been shown to be successfully grafted into animals, but there are still issues with complete vascularization, thrombosis, and integration with the host vasculature.

Kidney-on-a-chip: a microfluidic device mimicking the structure and function of human kidneys, which allows researchers to study renal biology, disease, and drug effects in vitro in a more accurate and scalable manner. Major advances include better mimicry of kidney microstructures; integration with other organ models; utility in drug testing, toxicity studies, and disease modeling—leading toward a promising development in personalized medicine.

Limitations

One major limitation in kidney engineering is the size and scale of the organ itself. The kidneys are relatively large and contain highly complex, densely branched networks of blood vessels that are essential for proper function. Replicating this intricate vascular architecture in a laboratory setting is extremely challenging, as a fully functional vascular network must be capable of supporting the entire organ. Even small imperfections in vessel structure or distribution can compromise oxygen delivery, nutrient exchange, and waste removal, making it difficult to sustain engineered kidney tissue over time.

Another significant challenge lies in vascular integration following transplantation. For an

engineered kidney to function in the body, its blood vessels must successfully connect with the host's circulatory system. Achieving stable anastomosis between grafted vessels and the recipient's vasculature without triggering thrombosis, immune rejection, or long-term vessel failure remains an area of ongoing research. The lack of complete understanding of how engineered vasculature behaves in vivo limits the reliability and safety of transplantation outcomes.

Finally, functional vascularization presents an additional layer of complexity. Blood vessels in the kidney must do more than simply transport blood; they play an active role in regulating blood pressure, delivering oxygen, and enabling precise filtration processes. Native kidneys rely on sophisticated feedback mechanisms to maintain proper flow and pressure within the glomeruli. Recreating these dynamic regulatory systems in engineered tissues has proven extremely difficult, as it requires precise control over vascular responsiveness and integration with renal filtration functions.

Recommendations

Hybrid strategies that combine biological scaffolding with advanced bioprinting techniques offer a promising solution to the persistent challenge of vascularization in kidney tissue engineering. For instance, integrating precisely 3D-printed vascular networks into decellularized kidney scaffolds may allow researchers to take advantage of the natural structural support and biochemical cues provided by the extracellular matrix while also benefiting from the spatial precision of bioprinted capillary systems. This combined approach has the potential to produce a more stable and functional vasculature capable of adequately supporting kidney tissue growth, perfusion, and long-term function.

The development of specialized bioinks represents another important avenue for improving vascular integration. Bioinks that incorporate both endothelial cells and kidney-specific cell types, such as proximal tubule cells, may enhance the endothelialization of 3D-printed vascular networks. By promoting closer interaction and communication between blood vessels and surrounding renal tissue, this strategy could significantly improve both structural integrity and functional recovery of engineered kidney constructs, which depend heavily on tight vascular–tissue coupling.

In addition, gene-editing technologies such as CRISPR-Cas9 present new opportunities to directly enhance vascularization at the cellular level. By modifying the genetic code of cells involved in angiogenesis, researchers may improve their ability to form organized and efficient blood vessel networks or accelerate tissue regeneration. For example, targeted editing of endothelial cells to increase their capacity for vascular network formation could serve as a directed and controlled strategy to enhance vascular development within engineered kidney tissues.

Finally, biomechanical stimulation is an emerging area of interest that may further improve the maturation and functionality of engineered kidney constructs. Native kidneys are constantly exposed to both passive and active mechanical forces, including shear stress from blood flow and cyclic stretching from filtration dynamics. Replicating these conditions in vitro through the use of bioreactors or organ-on-a-chip systems could promote better cellular organization and differentiation. When appropriately applied, mechanical stimuli have the potential to enhance the formation of functional renal structures, such as tubules and vascular networks, bringing engineered tissues closer to physiological reality.

Closing Thought

Chronic kidney disease (CKD) continues to represent a significant global health burden due to its late detection, irreversible damage at advanced stages, and steadily increasing incidence. Although lifestyle modifications and dietary interventions may help reduce risk, disease management is complicated by non-modifiable factors such as age, gender, and ethnicity. The limitations of kidney transplantation, including the shortage of donor organs and immune-related complications following surgery, highlight the need for alternative therapeutic approaches. As a result, regenerative medicine has emerged as a highly promising solution for addressing these challenges.

Despite its potential, kidney tissue engineering faces substantial barriers due to the structural and functional complexity of the kidney, particularly its intricate vascular network. Recent advances in techniques such as biological scaffolding, bioprinting, and kidney-on-a-chip models show considerable promise; however, they have not yet succeeded in fully replicating native kidney function, especially with respect to vascularization. Without adequate vascularization, engineered kidney tissues cannot sustain perfusion or perform essential renal functions, limiting the effectiveness of these approaches. As research in regenerative medicine continues to advance, efforts remain focused on developing effective and personalized treatments for CKD patients. In the long term, these strategies have the potential to reduce dependence on organ transplantation and improve patient quality of life, as has been achieved in regenerative approaches for other, less complex organs.

Works Cited

1. Ammirati, A. L. (2020). Chronic Kidney Disease. Revista Da Associação Médica Brasileira, 66(suppl 1), s03–s09. https://doi.org/10.1590/1806-9282.66.s1.3
2. Azhar, A., Hassan, N., Tapolyai, M., & Molnar, M. Z. (2021). Obesity, Chronic Kidney Disease, and Kidney Transplantation: An Evolving Relationship. Seminars in Nephrology, 41(2), 189–200. https://doi.org/10.1016/j.semnephrol.2021.03.013
3. Coresh, J., Selvin, E., Stevens, L. A., Manzi, J., Kusek, J. W., Eggers, P., Van Lente, F., & Levey, A. S. (2007). Prevalence of Chronic Kidney Disease in the United States. JAMA, 298(17), 2038. https://doi.org/10.1001/jama.298.17.2038
4. National Kidney Foundation. (2023, September 11). Chronic Kidney Disease (CKD). National Kidney Foundation. https://www.kidney.org/kidney-topics/chronic-kidney-disease-ckd
5. Ramírez, M. (2013). Multiple Organ Dysfunction Syndrome. Current Problems in Pediatric and Adolescent Health Care, 43(10), 273–277. https://doi.org/10.1016/j.cppeds.2013.10.003
6. Chapter 1: Definition and classification of CKD. (2013). Kidney International Supplements, 3(1), 19–62. https://doi.org/10.1038/kisup.2012.64
7. Vaidya, S., & Aeddula, N. (2024, July 31). Chronic kidney disease. Nih.gov; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK535404/
8. Bach, K. E., Kelly, J. T., Palmer, S. C., Khalesi, S., Strippoli, G. F. M., & Campbell, K. L. (2019). Healthy Dietary Patterns and Incidence of CKD. Clinical Journal of the American Society of Nephrology, 14(10), 1441–1449. https://doi.org/10.2215/cjn.00530119
9. Hu, E. A., Coresh, J., Anderson, C. A. M., Appel, L. J., Grams, M. E., Crews, D. C., Mills, K. T., He, J., Scialla, J., Rahman, M., Navaneethan, S. D., Lash, J. P., Ricardo, A. C., Feldman, H. I., Weir, M. R., Shou, H., Rebholz, C. M., Go, A. S., Rao, P. S., & Townsend, R. R. (2021). Adherence to Healthy Dietary Patterns and Risk of CKD Progression and All-Cause Mortality: Findings From the CRIC
(Chronic Renal Insufficiency Cohort) Study. American Journal of Kidney Diseases, 77(2), 235–244. https://doi.org/10.1053/j.ajkd.2020.04.019
10. Gulas, E., Wysiadecki, G., Szymański, J., Majos, A., Stefańczyk, L., Topol, M., & Polguj, M. (2018). Morphological and clinical aspects of the occurrence of accessory (multiple) renal arteries. Archives of Medical Science, 14(2), 442–453. https://doi.org/10.5114/aoms.2015.55203
11. Lo, D. J., Kaplan, B., & Kirk, A. D. (2014). Biomarkers for kidney transplant rejection. Nature Reviews Nephrology, 10(4), 215–225. https://doi.org/10.1038/nrneph.2013.281
12. Dulak, J., Szade, K., Szade, A., Nowak, W., & Józkowicz, A. (2015). Adult stem cells: hopes and hypes of regenerative medicine. Acta Biochimica Polonica, 62(3), 329–337.
https://doi.org/10.18388/abp.2015_1023
13. Arjmand, B., Rabbani, Z., Soveyzi, F., Tayanloo-Beik, A., Rezaei-Tavirani, M., Biglar, M., Adibi, H., & Larijani, B. (2022). Advancement of organoid technology in regenerative medicine. Regenerative Engineering and Translational Medicine, 9(1), 1–14.

https://doi.org/10.1007/s40883-022-00271-0
14. National Institute of Biomedical Imaging and Bioengineering. (2009). Tissue Engineering and Regenerative Medicine | National Institute of Biomedical Imaging and Bioengineering. Nih.gov; National Institute of Biomedical Imaging and Bioengineering. https://www.nibib.nih.gov/science-education/science-topics/tissue-engineering-and-regenerative-med icine
15. Soriano, R. M., & Leslie, S. W. (2023, July 24). Anatomy, Abdomen and Pelvis, Kidneys. Nih.gov; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK482385/
16. Shyam Venkatakrishna, Onyango, L. C., Serai, S. D., & Viteri, B. (2023). Kidney Anatomy and Physiology. Springer EBooks, 3–12. https://doi.org/10.1007/978-3-031-40169-5_1
17. Preuss, H. G. (1993). Basics of Renal Anatomy and Physiology. Clinics in Laboratory Medicine, 13(1), 1–11. https://doi.org/10.1016/s0272-2712(18)30456-6
18. Leslie, S. W., & Sajjad, H. (2023). Anatomy, Abdomen and Pelvis, Renal Artery. PubMed; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK459158/
19. Bohle, A., Christensen, J., Meyer, D. S., Laberke, H. G., & Strauch, M. (1982). Juxtaglomerular apparatus of the human kidney: Correlation between structure and function. | EBSCOhost. Ebsco.com, 12, S-18. https://doi.org/10.1111/1523-1755.ep17669467
20. Inglis, S., Christensen, D., Wilson, D. I., Kanczler, J. M., & Oreffo, R. O. C. (2016). Human endothelial and foetal femur-derived stem cell co-cultures modulate osteogenesis and angiogenesis. Stem Cell Research & Therapy, 7(1). https://doi.org/10.1186/s13287-015-0270-3
21. Figliuzzi, M., Remuzzi, G., & Remuzzi, A. (2014). Renal Bioengineering with Scaffolds Generated from Rat and Pig Kidneys. Nephron Experimental Nephrology, 126(2), 113–118. https://doi.org/10.1159/000360683
22. Sullivan, D. C., Mirmalek-Sani, S.-H., Deegan, D. B., Baptista, P. M., Aboushwareb, T., Atala, A., & Yoo, J. J. (2012). Decellularization methods of porcine kidneys for whole organ engineering using a high-throughput system. Biomaterials, 33(31), 7756–7764. https://doi.org/10.1016/j.biomaterials.2012.07.023
23. Katari, R., Peloso, A., Zambon, J. P., Soker, S., Stratta, R. J., Atala, A., & Orlando, G. (2014). Renal Bioengineering with Scaffolds Generated from Human Kidneys. Nephron Experimental Nephrology, 126(2), 119–124. https://doi.org/10.1159/000360684
24. Fransen, Maaike F. J., Addario, G., Bouten, Carlijn V. C., Halary, F., Moroni, L., & Mota, C. (2021). Bioprinting of kidney in vitro models: cells, biomaterials, and manufacturing techniques. Essays in Biochemistry, 65(3), 587–602. https://doi.org/10.1042/EBC20200158
25. Kačarević, Ž., Rider, P., Alkildani, S., Retnasingh, S., Smeets, R., Jung, O., Ivanišević, Z., & Barbeck, M. (2018). An Introduction to 3D Bioprinting: Possibilities, Challenges and Future Aspects. Materials, 11(11), 2199. https://doi.org/10.3390/ma11112199
26. Mota, C., Camarero-Espinosa, S., Baker, M. B., Wieringa, P., & Moroni, L. (2020).

Bioprinting: From Tissue and Organ Development to in Vitro Models. Chemical Reviews, 120(19), 10547–10607. https://doi.org/10.1021/acs.chemrev.9b00789
27. Meng, W., Wang Zhihai, Can, W., Fei, W., & Xi, C. (2018). Bioprinting. 201, 89–95. https://doi.org/10.1145/3208955.3208963
28. Park, T. Y., Yang, Y. J., Ha, D.-H., Cho, D.-W., & Cha, H. J. (2019). Marine-derived natural polymer-based bioprinting ink for biocompatible, durable, and controllable 3D constructs. Biofabrication, 11(3), 035001. https://doi.org/10.1088/1758-5090/ab0c6f
29. Yu, F., & Choudhury, D. (2019). Microfluidic bioprinting for organ-on-a-chip models. Drug Discovery Today, 24(6), 1248–1257. https://doi.org/10.1016/j.drudis.2019.03.025
30. Kim, S., & Takayama, S. (2015). Organ-on-a-chip and the kidney. Kidney Research and Clinical Practice, 34(3), 165–169. https://doi.org/10.1016/j.krcp.2015.08.001
31. Bhatia, S. N., & Ingber, D. E. (2014). Microfluidic organs-on-chips. Nature Biotechnology, 32(8), 760–772. https://doi.org/10.1038/nbt.2989
32. Kimura, H., Yamamoto, T., Sakai, H., Sakai, Y., & Fujii, T. (2008). An integrated microfluidic system for long-term perfusion culture and on-line monitoring of intestinal tissue models. Lab on a Chip, 8(5), 741. https://doi.org/10.1039/b717091b
33. Huh, D., Matthews, B. D., Mammoto, A., Montoya-Zavala, M., Hsin, H. Y., & Ingber, D. E. (2010). Reconstituting organ-level lung functions on a chip. Science (New York, N.Y.), 328(5986), 1662–1668. https://doi.org/10.1126/science.1188302
34. Franco, C., & Gerhardt, H. (2012). Blood vessels on a chip. Nature, 488(7412), 465–466. https://doi.org/10.1038/488465a
35. Zhang, Y. S., Zhang, Y.-N., & Zhang, W. (2017). Cancer-on-a-chip systems at the frontier of nanomedicine. Drug Discovery Today, 22(9), 1392–1399.
https://doi.org/10.1016/j.drudis.2017.03.011
36. Kefallinou, D., Grigoriou, M., Boumpas, D. T., Gogolides, E., & Tserepi, A. (2020). Fabrication of a 3D microfluidic cell culture device for bone marrow-on-a-chip. Micro and Nano Engineering, 9, 100075. https://doi.org/10.1016/j.mne.2020.100075
37. Zanetti, F. (2020, January 1). Chapter 7 - Kidney-on-a-chip (J. Hoeng, D. Bovard, & M. C. Peitsch, Eds.). ScienceDirect; Academic Press.
https://www.sciencedirect.com/science/article/abs/pii/B9780128172025000073
38. Yurchenco, P. D. (2010). Basement Membranes: Cell Scaffoldings and Signaling Platforms. Cold Spring Harbor Perspectives in Biology, 3(2), a004911–a004911.
https://doi.org/10.1101/cshperspect.a004911
39. Wang, D., Gust, M., & Ferrell, N. (2022). Kidney-on-a-Chip: Mechanical Stimulation and Sensor Integration. Sensors, 22(18), 6889. https://doi.org/10.3390/s22186889
40. Jang, K.-J., & Suh, K.-Y. (2010). A multi-layer microfluidic device for efficient culture and analysis of renal tubular cells. Lab Chip, 10(1), 36–42. https://doi.org/10.1039/b907515a
41. Jang, K.-J., Mehr, A. P., Hamilton, G. A., McPartlin, L. A., Chung, S., Suh, K.-Y., & Ingber, D. E. (2013). Human kidney proximal tubule-on-a-chip for drug transport and nephrotoxicity assessment. Integrative Biology, 5(9), 1119–1129.

https://doi.org/10.1039/c3ib40049b
42. Ashammakhi, N., Wesseling-Perry, K., Hasan, A., Elkhammas, E., & Zhang, Y. S. (2018). Kidney-on-a-chip: untapped opportunities. Kidney International, 94(6), 1073–1086. https://doi.org/10.1016/j.kint.2018.06.034
43. Aazmi, A., Zhou, H., Li, Y., Yu, M., Xu, X., Wu, Y., Ma, L., Zhang, B., & Yang, H. (2022). Engineered Vasculature for Organ-on-a-Chip Systems. Engineering, 9, 131–147. https://doi.org/10.1016/j.eng.2021.06.020
44. Rayner, S. G., Phong, K. T., Xue, J., Lih, D., Shankland, S. J., Kelly, E. J., Himmelfarb, J., & Zheng, Y. (2018). Reconstructing the Human Renal Vascular–Tubular Unit In Vitro. Advanced Healthcare Materials, 7(23), 1801120. https://doi.org/10.1002/adhm.201801120
45. Homan, K. A., Gupta, N., Kroll, K. T., Kolesky, D. B., Skylar-Scott, M., Miyoshi, T., Mau, D., Valerius, M. T., Ferrante, T., Bonventre, J. V., Lewis, J. A., & Morizane, R. (2019). Flow-enhanced vascularization and maturation of kidney organoids in vitro. Nature Methods, 16(3), 255–262. https://doi.org/10.1038/s41592-019-0325-y
46. Sánchez, P. F., Brey, E. M., & Briceño, J. C. (2018b). Endothelialization mechanisms in vascular grafts. Journal of Tissue Engineering and Regenerative Medicine, 12(11), 2164–2178. https://doi.org/10.1002/term.2747
47. Wu, W., Duan, Y., Ma, G., Zhou, G., Park-Windhol, C., D'Amore, P. A., & Lei, H. (2017). AAV-CRISPR/Cas9–Mediated Depletion of VEGFR2 Blocks Angiogenesis In Vitro. Investigative Opthalmology & Visual Science, 58(14), 6082. https://doi.org/10.1167/iovs.17-21902

A Study of Sentiment Detection in Tweets Using Transformer Embeddings and Standard Classifiers

Kerem Tuzel
keremtuzel.kt@gmail.com

Abstract

Detecting emotion in brief, informal text is a challenging task in natural language processing, further complicated by mixed emotions and irregular grammar. This study evaluates several machine learning models for multi-class emotion detection in tweets. Using 40,000 X tweets labeled with 13 emotions we tested four models including Decision Tree, Random Forest, XGBoost, and Multilayer Perceptron (MLP). We encoded each tweet using an all-MiniLM-L12-v2 transformer into 384-dimensional vectors. Models were trained on 90% of the data and tested on the remaining 10%. Performance was determined via accuracy and weighted/macro F1 scores. All models exceeded the random guess baseline of 1 in 13 (7.7%) with XGBoost and MLP achieving the best performance (accuracy of 36%, weighted f1 of 0.32). While frequent emotions such as Neutral, Worry and Happiness were predicted with high accuracy, rare emotions were misclassified. These findings highlight both the potential and limitations of current approaches for granular emotion detection and point to the importance of addressing class imbalance and domain specific language modeling in future work.

1 Introduction

When artificial intelligence (AI) is discussed, it is often imagined as a machine designed to complete human tasks, from data processing to decision making. One field where AI is perceived to struggle in is recognizing and understanding of human emotion, a topic considered too nuanced and complicated for machines to accurately interpret. However, given the significant advances in machine learning (ML), natural language processing (NLP) [1] and deep learning, it is plausible to train models that can detect and classify such subjective and complex phenomena of human sentiments in written text[2].

This paper investigates the ability of NLP [1] and AI models to detect emotions in informal online text, specifically tweets. As social media becomes a primary outlet for emotional expression, accurate sentiment classification can prove valuable in multiple applications, including marketing, customer service, and public opinion monitoring. In this paper, we aim to build a model that can identify one of 13 distinct emotions: empty, sadness, enthusiasm, neutral, worry, surprise, love, fun, hate, happiness, boredom, relief, and anger.

Emotion plays a crucial role in human communication. The tone of a sentence can dramatically change its meaning and how it is interpreted. Sarcasm, irony, and hidden frustration can make sentiment detection especially challenging. Without robust emotion recognition, AI risks misinterpreting user intention, potentially failing to detect mental health distress, overlooking customer dissatisfaction, or misreading public sentiment. As AI becomes increasingly integrated into our daily lives, the ability to detect and understand emotions accurately becomes an essential goal rather than a desirable capability.

NLP focuses on enabling machines to both understand and generate human language. The introduction of transformer-based architectures such as Bidirectional Encoder Representations from Transformers (BERT) [3], Robustly Optimized BERT Pretraining Approach (RoBERTa) [4], and Generative Pre-trained Transformer (GPT) [5] has greatly advanced the field. These models leverage deep neural networks and attention mechanisms to capture contextual relationships between words, making them especially effective for texts requiring subjective and nuanced language understanding.

We hypothesize that, with an appropriate dataset, it is possible to train models capable of labeling emotions in tweets with meaningful accuracy. This study compares Decision Trees [6], Random Forests [7], XGBoost [8], and Multilayer Perceptron (MLP) [9], and explores an ensemble approach that leverages the strengths of multiple models.

2 Related Work

Nandwani and Verma provide a comprehensive review of sentiment analysis and emotion detection methods, comparing lexicon-based, machine learning, and deep learning approaches [10]. Their study finds that deep learning models, particularly LSTMs and RNNs, perform well on large datasets, while lexicon-based and traditional ML methods are improving due to better feature engineering. Key challenges include sarcasm, mixed emotions, and informal social media text such as new slang, spelling errors and misuse of grammar.

Abas et al. propose a hybrid deep learning model, BERT-CNN, combining BERT's contextual embeddings with CNN classification layers for emotion detection [11]. Tested on the SemEval 2019 Task 3 and ISEAR datasets, their model achieved 94.7% accuracy and a 94% F1-score on SemEval, outperforming existing models.

Machoba et al. explore a machine learning approach combining lexicon-based methods, traditional classifiers, and deep learning [12]. Their hybrid CNN–RNN model achieved $\tilde{9}0\%$ accuracy across six basic emotions, but struggled with sarcasm, irony, and overlapping emotions.

Qi and Shabrina use X data from major cities in England to study public sentiment toward Covid-19 during the third national lockdown as a case study to evaluate lexicon-based and supervised machine learning methods [13]. Support Vector Classifier (SVC) models with Bag of Words (BoW) and Term Frequency–Inverse Document Frequency (TF–IDF) features achieve the highest accuracy at 71%, with trends showing public opinion in an initial rise in positivity which was followed by a decline.

Ghatora et al. compare traditional machine learning models with pre-trained Large Language Models (LLMs) for sentiment analysis of Flipkart product reviews [14]. Support Vector Machine (SVM) performs best for shorter, more concise reviews, while GPT-4 excels on longer, context filled text by capturing nuanced sentiments and providing detailed explanations. The authors note the potential for integrating LLMs into business strategies to enhance product feedback analysis and suggest future exploration of open-source models and efficiency improvements.

Chakriswaran et al. review emotion AI-driven sentiment analysis methods, researching ontology-based, lexicon-based, machine learning, and neural network approaches across multiple domains [15]. On sample X data, the aspect-based ontology method achieved 83% accuracy, the term frequency approach achieved 85%, and an SVM-based model 90%. The authors emphasize strengths of traditional methods and suggest incorporating deep learning to further improve performance.

3 Materials and Methods

3.1 Dataset

The dataset, sourced from Kaggle's Emotion Detection from Text competition (2021) [16], contains 40,000 tweets labeled with one of 13 emotions (Figure 1). Each piece of data has a tweet_id, a sentiment label, and content of the tweet. The entire dataset was manually edited by humans, making the sentiment field very precise. Duplicate tweets were removed to ensure unique examples.

	tweet_id	sentiment	content
0	1956967341	empty	i know i was listenin to bad habi...
1	1956967666	sadness	Layin n bed with a headache ughhhh...waitin o...
2	1956967696	sadness	Funeral ceremony...gloomy friday...
3	1956967789	enthusiasm	wants to hang out with friends SOON!
4	1956968416	neutral	We want to trade with someone w...

Table 1: Example data showing tweet_id, sentiment label, and the content of the tweet in each data piece.

3.2 Preprocessing and Encoding

Tweets and corresponding sentiment were all in raw text format. We encoded these texts using the all-MiniLM-L12-v2 [17] sentence transformer, producing 384-dimensional vectors via WordPiece tokenization, positional and segment embeddings, and a 12-layer transformer architecture with mean pooling [18]. In order to preserve the sentence structure and word order, the transformer has a corresponding learned positional embedding, for each token's position, which is also 384 dimensions. If the input is one sentence,

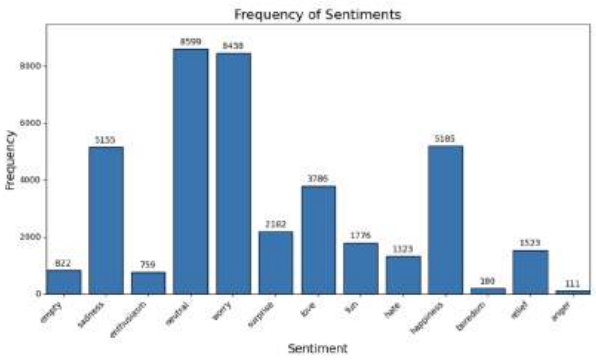


Figure 1: Frequency of each emotion throughout the dataset seen in a histogram.

segment embeddings are still included, but simply set to zero, which allows it to remain compatible with BERT's [3] input formatting. The sum of the three embeddings (token, positional, and segment), form the initial embedding which is then processed through the 12-layer MiniLM transformer, where multi-head, self-attention, and feed-forward networks embed information from throughout the sequence. This allows each token vector to include not only its own meaning but also how it relates to the other tokens within the context. Finally, mean pooling is applied over all valid token vectors to produce a single 384-dimensional embedding for each tweet.

Emotions were mapped to integers (0–12) corresponding with sentiments: empty, sadness, enthusiasm, neutral, worry, surprise, love, fun, hate, happiness, boredom, relief, and anger, respectively. Data was split into 90% training and 10% testing sets using scikit-learn's train_test_split, with a fixed random seed for reproducibility.

3.3 Models

We evaluated four different machine learning algorithms for emotion classification: Decision Tree [6], Random Forest [7], XGBoost [8], and MLP [9]. Each model was trained using the same 384-dimensional MiniLM [17] sentence embeddings as input and configured through parameter tuning to maximize performance.

Decision Tree

Implemented using scikit-learn's DecisionTreeClassifier. Decision trees [6] make predictions by recursively splitting the feature space based on learned thresholds, creating internal nodes until reaching leaf nodes that assign class labels. The primary hyperparameter, max_depth, was tested from 2 to 20 in increments of 2 to balance underfitting and overfitting. A Decision Tree [6] is a composition of layers which each have their own number of nodes. At the top layer there is one node which the input is given to. At each node, a decision is made and the input is affected and given to another node on the next layer, till at the bottom the final output should come out as a valid number output. The max_depth is the maximum number of layers the tree can grow from the root node to a leaf node.

Random Forest

Random forests [7] are ensembles of Decision Trees [6], each trained on bootstrap samples of the data and considering a random subset of features at each split to reduce correlation between trees. Using scikit-learn's RandomForestClassifier, we tested max_depth values of 5, 10, 15, and 20, and n_estimators of 50, 100, 150, and 200. Final predictions were made via majority voting across trees. The max_depth is the maximum number of layers each tree can grow from the root node to a leaf node and the n_estimators determines the amount of individual trees.

XGBoost

XGBoost [8] is an optimized gradient boosting framework that builds trees sequentially, with each new tree correcting errors from the previous ensemble. We tuned max_depth=4, n_estimators=400, learning_rate=0.1, subsample=1.0, and colsample_bytree=0.1. The multi:softprob objective was used for probability outputs, with mlogloss as the evaluation metric.

Multilayer Perceptron (MLP)

The MLP [9] is a feed-forward neural network capable of modeling complex nonlinear relationships. Using scikit-learn's MLPClassifier, we tested max_iter values from 10 to 3000 to find the smallest number of iterations that achieved the highest validation accuracy. The max_depth is the maximum number of layers each tree can grow from the root node to a leaf node. The n_estimators determines the number of individual trees in the ensemble. The learning_rate scales the contribution of each tree to the final model, controlling how quickly the model adapts. The

subsample divides the training data used for each tree. The colsample_bytree divides the features randomly sampled for each tree, helping increase diversity among trees.

Weighted Majority Ensemble

We combined models via weighted majority voting, starting with equal weights (1.0) and halving a model's weight for each incorrect prediction. Throughout this process, the models which perform more poorly would have their weight or "opinion" greatly diminished, while the more successful ones would come out having the value of their predictions increased. The class with the highest weighted vote sum was chosen as the final prediction.

4 Results

Across all models tested, performance exceeded the random-guess baseline of 7.7% accuracy (1 in 13 for each sentiment), confirming that transformer-based embeddings capture meaningful semantic cues for emotion detection in informal text. Among the models, XGBoost [8] and MLP [9] achieved the highest performance, each reaching an accuracy of 36% and a weighted F1-score of 0.32 (Table 2)

Model	Best Parameters	Accuracy	Weighted Avg F1 Score
Decision Tree	max_depth: 4	0.26	0.19
Random Forest	max_depth: 20, num_estimators: 200	0.32	0.25
X.G. Boost	max_depth: 4, num_estimators: 400, learning_rate = 0.1, subsample: 1.0, colsample_bytree: 0.1	0.36	0.32
MLP	max_iter: 20	0.36	0.32

Table 2: Best parameters for each of the four models, as well as their achieved accuracies and weighted average f1 scores.

Frequent emotions such as Neutral, Worry, and happiness were predicted with moderate success, often achieving recall values above 0.4 (Table 3). In contrast, rare emotions such as, Empty, Boredom, and Anger had near-zero recall, indicating they were not correctly identified. This imbalance in performance per class reflects the skewed distribution of the dataset, where certain emotions dominate the training data.

	Decision Tree		Random Forest		XGBoost		MLP	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
Empty	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sadness	0.00	0.00	0.41	0.07	0.35	0.25	0.35	0.33
Enthusiasm	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Neutral	0.25	0.54	0.30	0.61	0.37	0.58	0.36	0.56
Worry	0.28	0.41	0.31	0.58	0.35	0.50	0.36	0.42
Surprise	0.00	0.00	0.00	0.00	0.13	0.01	0.20	0.03
Love	0.58	0.10	0.52	0.23	0.45	0.36	0.46	0.40
Fun	0.00	0.00	0.00	0.00	0.21	0.02	0.26	0.02
Hate	0.00	0.00	0.00	0.00	0.33	0.09	0.31	0.16
Happiness	0.22	0.35	0.36	0.24	0.33	0.40	0.32	0.45
Boredom	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Relief	0.00	0.00	0.00	0.00	0.25	0.03	0.15	0.02
Anger	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 3: Precision and Recall of each model when predicting each sentiment in the testing set.

Per Class Trends:

- **High Recall**: Neutral, Worry, and Happiness showed highest recall due to class frequency.
- **Moderate Recall**: Love, Hate, and Surprise had some detection ability, but prone to confusion with related classes.
- **Low Recall**: Empty, Boredom, Fun, and Anger were rarely detected due to limited data.

Figure 2 (Confusion matrix) further highlights this pattern; strong diagonal elements for majority classes contrast with sparse correct predictions for minority classes.

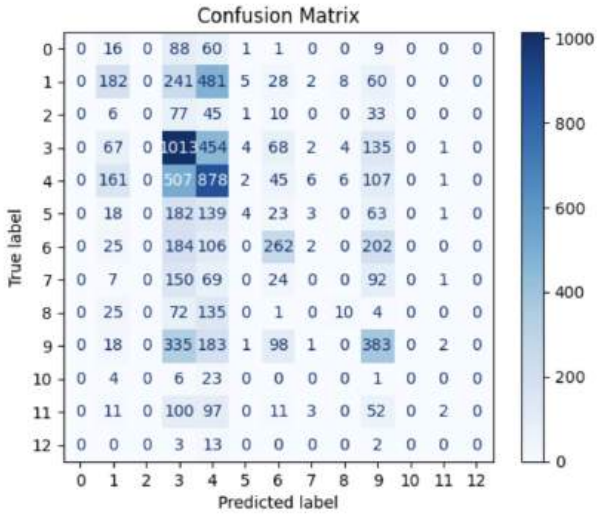


Figure 2: Confusion Matrix of performance of XGBoost model. The number in each cell shows how many times the column label was predicted for a tweet with the true label in the row. Sentiments: empty, sadness, enthusiasm, neutral, worry, surprise, love, fun, hate, happiness, boredom, relief, and anger corresponded to values: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12, respectively.

Precision-Recall Example: Neutral Tweets

To illustrate how precision and recall are manifested in our model, Figure 3 shows the classification break down for the Neutral emotion:

- True Positives: correctly predicted Neutral tweets.
- False Positives: tweets predicted as Neutral but belonging to another class.
- False Negatives: Actual Neutral tweets predicted as another emotion.

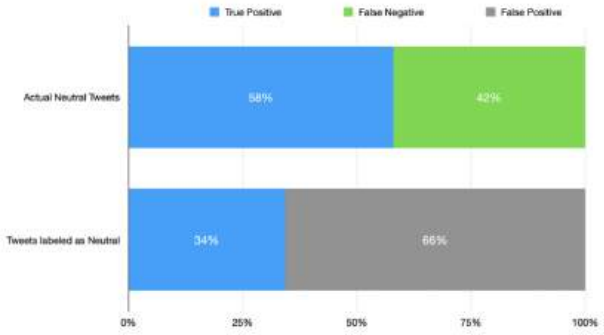


Figure 3: Comparison of false positives, false negatives, and true positives for the neutral sentiment from the XGBoost model.

This example reveals that, while Neutral has relatively high recall, meaning most actual Neutral tweets were found, the precision is lower due to substantial false positives, often from tweets labeled as Worry or Happiness.

5 Discussion

The results demonstrate that all four models exceeded the random baseline. XGBoost [8] and MLP [9] achieved the highest performance, indicating both neural networks and gradient-boosted trees can effectively leverage dense sentence embeddings. The MLP [9] likely benefitted from its ability to capture nonlinear patterns in dense embeddings, while XGBoost [8] leveraged gradient boosting to focus on challenging examples. Decision Trees [6] tended to overfit with deeper depths, whereas Random Forests [7] improved with more estimators but reached a level of stability in accuracy.

Performance skew toward frequent classes reflects the well-documented challenge of class imbalance in sentiment analysis. Frequent emotions dominate model predictions, while rare emotions are often ignored. This was heavily seen when looking into the true positives, false positives, and false negatives of the Neutral class in XGBoost [8]. The model was able to correctly identify most of the Neutral tweets, but also had a very large amount of false positives which meant it also mislabeled a lot of tweets as Neutral, implying that it had favored that class greatly when choosing a sentiment. The false positives largely come from tweets whose true labels are Worry or Happiness. This pattern is consistent with the class distribution in the dataset: Neutral, Worry, and Happi-

ness are among the most frequent emotions, and their linguistic expressions often overlap in tone and content, particularly in informal X text where subtle emotional cues may be sparse or ambiguous. In an imbalanced dataset, models are naturally biased toward predicting frequent classes, both because they dominate during training and because misclassifying between them results in a smaller penalty during optimization. Since these classes are so prevalent in the training data, the model learns strong decision boundaries for them and leans towards them by default when uncertainty is high. As a result, tweets with ambiguous or weakly expressed emotions are more likely to be labeled as one of these high-frequency categories, explaining the inflation of false positives for Neutral while maintaining high recall. On the other side of the spectrum, this lead to underrepresented classes in the dataset to be guessed at a very low rate, some being never guessed at all. This limitation is consistent with prior studies, including Nandwani & Verma's [10] finding that informal, imbalanced datasets hinder rare emotion detection.

Additionally, the general-purpose MiniLM embeddings, while effective overall, were not fine-tuned for X linguistic style [18]. Texts on social media rarely follow the formal structure of language that is used to pre-train such models. Slang, emojis, abbreviations, and sarcasm may be underrepresented in the learned space, or entirely absent. This reduces model sensitivity to certain sentiment cues, especially on a platform where emotional expression often relies on subtle hints of sarcasm and irony. Without fine-tuning on a large, representative sample of X content, the embeddings may map these features less accurately, leading models to overlook or misinterpret certain emotional signals. This flaw likely compounds the issues that come with class imbalance, as rare emotions that depend heavily on such nuanced details become even harder for the model to detect reliably, or at all.

Regarding future works, we plan to address the current imbalance and performance skew toward frequent classes through both data and model-level interventions. At the data level, techniques such as oversampling and SMOTE adapted for dense embeddings could be employed, alongside targeted collection of tweets which are severely underrepresented in the data set. For the models, incorporating class weighting into training, fine-tuning transformer embeddings specifically on X data, and exploring hybrid architectures (e.g., BERT + CNN) [11] may enhance sensitivity to subtle cues. Finally, in evaluating future models, we will prioritize macro F1-score during hyperparameter tuning to encourage balanced performance across all sentiment classes, rather than favoring frequent labels that show up frequently.

6 Conclusion

This study demonstrates that transformer-based embeddings paired with standard classifiers can substantially improve fine-grained emotion detection in tweets over random guessing. XGBoost [8] and MLP [9] yielded the best results (accuracy = 0.36, weighted F1 = 0.32), but performance was limited by extreme class imbalance and the use of embeddings not specialized for informal, emotion-rich language. Sarcasm remains a particular challenge, functioning as a modifier that can invert the sentiment of otherwise sincere expressions, completely changing the meaning of the sentence. Reliable detection therefore requires approaches that are able to identify such modifiers as well as the underlying emotion. Addressing these issues through balanced data augmentation and domain-specific model fine-tuning offers the most promising path toward more equitable and accurate emotion detection.

References

[1] Kang, Yue, et al. "Natural Language Processing (NLP) in Management Research: A Literature Review." *Journal of Management Analytics*, vol. 7, no. 2, 2 Apr. 2020, pp. 139–172, https://doi.org/10.1080/23270012.2020.1756939.

[2] Dang, Nhan Cach, et al. "Sentiment Analysis Based on Deep Learning: A Comparative Study." *Electronics*, vol. 9, no. 3, 14 Mar. 2020, p. 483, https://doi.org/10.3390/electronics9030483.

[3] Devlin, Jacob, et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *arXiv preprint*, 24 May 2019. https://doi.org/10.48550/arXiv.1810.04805.

[4] Liu, Yinhan, et al. "Roberta: A robustly optimized bert pretraining approach." *arXiv*

preprint, 26 Jul. 2019, https://doi.org/10.48550/arXiv.1907.11692.

[5] Radford, Alec, et al. "Improving language understanding by generative pre-training.", 10 Sep. 202, p. 3.

[6] Priyanka, N.A., and Dharmender Kumar. "Decision Tree Classifier: A Detailed Survey." *International Journal of Information and Decision Sciences*, vol. 12, no. 3, 30 Apr. 2020, p. 246, https://doi.org/10.1504/ijids.2020.108141.

[7] Breiman, Leo. "Random Forests." *Machine Learning*, vol. 45, no. 1, Oct. 2001, pp. 5–32, https://doi.org/10.1023/a:1010933404324.

[8] Chen, Tianqi, and Carlos Guestrin. "XGBoost." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13 Aug. 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.

[9] Popescu, Marius-Constantin, et al. "Multilayer perceptron and neural networks." *WSEAS Transactions on Circuits and Systems* 8.7 July 2009, pp. 579-588.

[10] Nandwani, Pansy, and Rupali Verma. "A Review on Sentiment Analysis and Emotion Detection from Text." *Social Network Analysis and Mining*, vol. 11, no. 1, 28 Aug. 2021, https://doi.org/10.1007/s13278-021-00776-6.

[11] R. Abas, Ahmed, et al. "Bert-CNN: A Deep Learning Model for Detecting Emotions from Text." *Computers, Materials & Continua*, vol. 71, no. 2, 2022, pp. 2943–2961, https://doi.org/10.32604/cmc.2022.021671.

[12] Machová, Kristína, et al. "Detection of Emotion by Text Analysis Using Machine Learning." *Frontiers in Psychology*, vol. 14, 20 Sept. 2023, https://doi.org/10.3389/fpsyg.2023.1190326.

[13] Qi, Yuxing, and Zahratu Shabrina. "Sentiment Analysis Using Twitter Data: A Comparative Application of Lexicon- and Machine-Learning-Based Approach." *Social Network Analysis and Mining*, vol. 13, no. 1, 9 Feb. 2023, https://doi.org/10.1007/s13278-023-01030-x.

[14] Ghatora, Pawanjit Singh, et al. "Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM." *Big Data and Cognitive Computing*, vol. 8, no. 12, 23 Dec. 2024, p. 199, https://doi.org/10.3390/bdcc8120199.

[15] Chakriswaran, Priya, et al. "Emotion AI-Driven Sentiment Analysis: A Survey, Future Research Directions, and Open Issues." *Applied Sciences*, vol. 9, no. 24, 12 Dec. 2019, p. 5462, https://doi.org/10.3390/app9245462.

[16] "Emotion Detection from Text", *Kaggle Public Dataset*, https://www.kaggle.com/datasets/pashupatigupta/emotion-detection-from-text

[17] Reimers, Nils, and Iryna Gurevych. *all-MiniLM-L12-v2*. Hugging Face, https://huggingface.co/sentence-transformers/all-MiniLM-L12-v2.

[18] Wang, Liya, et al. "Adapting Sentence Transformers for the Aviation Domain." *AIAA SCITECH 2024 Forum*, 4 Jan. 2024, https://doi.org/10.2514/6.2024-2702.

Nutrition and Wellness in Pediatric Oncology

AP Capstone Research

April 2025

Word Count: 4,401

Abstract

The goal of this project was to determine to what extent nutrition impacts the health and mental wellness of pediatric oncology patients. Using a mixed questionnaire approach, composed of a questionnaire distributed to healthcare professionals, I gathered sufficient data to answer my research question. My new understanding highlights the critical role of nutrition in improving both physical health and psychological well-being during cancer treatment. This finding confirms the hypothesis I established during the initial phase of the research process. The study reveals that proper nutrition significantly contributes to better treatment outcomes, while poor nutrition can heighten stress levels and hinder physical recovery. It addresses a key gap in pediatric oncology research by examining the dual impact of nutrition on both physical and mental wellness in children undergoing cancer treatment. These findings suggest the need for healthcare professionals including doctors, nurses, and dietitians to adopt a nuanced approach that combines nutritional and psychological support in treatment plans. Additionally, the research emphasizes the importance of involving families in care, enabling them to advocate for comprehensive, patient-centered strategies that address both physical and emotional needs. The study also highlights opportunities for future research, including the inclusion of direct input from patients and families, the use of longitudinal designs to track outcomes over time, and the development of targeted interventions that support both nutrition and mental health. Overall, this research illustrates the essential role of nutrition in the holistic treatment of pediatric oncology patients.

Keywords: pediatric oncology, mental wellness, patient outcomes, psychological support, nutrition

Introduction

"The connotation of the word cancer in most languages includes the threat to life and the association of death and dying" (Muthny, Koch, & Stump, 1990). Cancer is defined as the uncontrollable growth and division of abnormal cells that invade neighboring parts of the body and spread to other organs, this process is known as metastasizing (World Health Organization [WHO], 2021). Tragically, "Each year an estimated 400,000 children and adolescents of 0-19 years old develop cancer" (World Health Organization [WHO], 2021), making it one of the leading causes of death among children globally (Duffy, 1953).

The standard treatment after a cancer diagnosis is composed of surgery, chemotherapy or radiotherapy and occasionally a combination of the three (World Health Organization [WHO], 2021).Despite advances in treatment, factors such as delayed diagnosis, abandonment of treatment, toxicity-related deaths, and poor nutrition continue to lower survival rates for pediatric oncology patients (World Health Organization [WHO], 2021).

Nutrition has long been recognized as a crucial factor in cancer care, influencing both prevention and recovery. Studies have shown that a poor diet can increase the likelihood of developing cancer, while dietary changes can reduce cancer risk by one to two thirds (Trichopoulos & Willett, 1996; Austoker, 1994). This is especially important for children, whose nutritional needs are closely tied to cognitive development and emotional stability. "Even short-term nutritional deficiencies can influence children's behavior, ability to concentrate, and perform complex tasks" (Young Children, 1998). Furthermore, poor nutrition has been linked to higher levels of chronic stress among adolescents (Thoresen & Eagleston, 1983).

Given these connections, it's clear that nutrition has a profound impact not only on the physical health but also on the mental wellness of pediatric oncology patients. This insight led

me to the research question: To what extent does nutrition impact the health and mental wellness of pediatric oncology patients?

Literature Review

Pediatric Oncology

Pediatric oncology refers to cancers that affect children aged zero to eighteen. A variety of cancers that impact pediatric patients with the most common sites being bone, blood and lymphatic tissue, soft somatic structures, kidney, eye and its orbit and the nervous system (Duffy, 1953). The main types of bone cancers in children are osteosarcoma and endothelioma while cancers of the blood manifest as leukemia or cancer of the lymphatic system. Soft tissue tumors are typically sarcomas. One of the more specific cancers of the kidney in children Wilms tumor and eye tumors in this age range are characterized by gliomas of the retina (Duffy, 1953). Additionally, brain tumors are more common in children than adults, making cancers of the nervous system a significant concern (Duffy, 1953).

A cancer diagnosis can be detrimental to a child's livelihood impacting their quality of life and resulting in a number of pervasive symptoms (Muthny, Koch, & Stump, 1990). In fact, a study investigating the side effects of cancer treatment through surveying people with bronchial carcinoma (after surgical operation), leukemia (CML and CLL), lymphoma (Hodgkin and non-Hodgkin lymphoma) as well as colorectal cancer found a third of them were dissatisfied with their physical functioning (Muthny, Koch, & Stump, 1990). Further interviews researchers conducted with with pediatric oncology patients, diagnosed with a variety of diagnoses including such as Acute lymphoblastic leukemia, brain tumors, lymphoma, osteosarcoma demonstrate that pediatric patients recognize the severity of their illnesses and the impact it has on their wellbeing (Hinds et al., 2004). One patient, when asked, "How has being sick been for you?" responded, "I

would tell you that I wonder why I am the one to get leukemia, but I tell myself it is better that I get it than my little sister" (Hinds et al., 2004).

The pre-existing research indicates that pediatric oncology patients are significantly impacted by their diagnosis. Therefore, it is of great importance to conduct further research to alleviate these severe symptoms and investigate what role nutrition plays in their health and mental wellness.

Impact of Nutrition on Children

Nutrition has proven to play a pivotal role in the health and wellness of children (Nutrition Of School Children, 1935). Inadequate nutrition among adolescents has been linked to increased fatigue, apathy and difficulties in forming relationships and engaging with their surroundings (Young Children, 1998). Additionally, research has proven that nutritional deficiencies result in lower attention span and in severe cases permanent cognitive damage (Young children, 1998).

Adequate nutrition during early childhood is essential for proper growth and development. It is important to educate children on healthy eating habits, as these behaviors often extend into adulthood and can help prevent chronic diseases (Herr & Morse, 1982). For instance, increasing fruit and vegetable intake has been associated with a reduced risk of diseases such as ischemic heart disease, stroke, and cancers of the lung, stomach, colon, and esophagus (Lock, Pomerleau, Causer & McKee, 2004). Based on a global review of dietary data and systematic meta-analyses, it estimates that increasing fruit and vegetable consumption could significantly reduce disease burdens, up to 31% for ischemic heart disease and 19% for stomach cancer (Lock, Pomerleau, Causer & McKee, 2004). Based on a global review of dietary data and systematic meta-analyses, increased consumption of fruits and vegetables could reduce the

disease burden by up to 31% for ischemic heart disease and 19% for stomach cancer (Lock, Pomerleau, Causer & McKee, 2004).

Overall, the existing research suggests that a nutrient-rich diet benefits children and adolescents by improving overall health, supporting development, and reducing the risk of disease.

Mental Wellness of Children and Adolescents

There are a number of mental disorders that impact children and adolescents such as ADHD and autism (Cuellar, 2015). ADHD is the most common mental disorder among young children, while depression is more prevalent among teenagers (Cuellar, 2015). Cuellar (2015) also emphasizes that "No single system in the U.S. identifies and treats children with mental disorders" (p. 116). Research has further shown that "Child mental health problems have long-term negative consequences, including lower educational attainment, lower wages, lower likelihood of employment, and more crime. Moreover, the negative impact of early mental health problems persists even if mental health later improves" (Cuellar, 2015, p. 112). This highlights the importance of early prevention, identification, and treatment.

Additionally, poverty has a significant impact on children's mental health. McLeod and Shanahan (1993) found that persistent poverty significantly predicts internalizing symptoms for instance, anxiety and depression, while current poverty is more strongly linked to externalizing behaviors like aggression (McLeod & Shanahan, 1993).

Many children and adolescents also face chronic stress. Chronic stress, is defined as a persistent imbalance between life demands and available resources, which has detrimental effects on physical and mental health especially in young people (Thoresen & Eagleston, 1983).

Thoresen and Eagleston (1983) warn that “Too much stress, however, destroys character, promoting a number of unhealthy and, in some cases, antisocial ways of miscoping” (p. 55).

The existing research solidifies that children and teenagers commonly experience chronic stress, ADHD, and depression. Therefore, their mental health must be prioritized, particularly when they are undergoing life-altering events such as a cancer diagnosis.

Gap in the Research

While existing research has explored the general impact of nutrition on patient health outcomes, there is limited investigation specifically addressing how nutrition influences both physical health and mental wellness in pediatric oncology patients. This population faces unique nutritional needs and mental health challenges due to the aggressive nature of cancer treatments, which can impair growth, immunity, and emotional resilience. Yet, the intersection of these two factors, nutrition and mental wellness, in this vulnerable group remains underexplored.

My research addresses this gap by focusing on the dual impact of nutrition on both the physical health and mental wellness of pediatric oncology patients, aiming to provide insights that could guide more holistic treatment protocols for this population.

Hypothesis and Assumptions

I hypothesize that healthcare professionals will recognize nutrition as a key factor in both the physical health and mental wellness of pediatric oncology patients. Additionally, I expect them to emphasize how nutrition influences children's overall well-being, particularly during cancer treatment.

Research Design and Methodology

Study Design

This study aims to explore how nutrition influences both the physical health and mental well-being of pediatric oncology patients. The goal is to address children with cancer, a previously overlooked group facing distinct nutritional needs and mental health challenges resulting from the aggressive nature of cancer treatments. Prior research highlights the importance of nutrition for children in general, but overlooks their mental wellness and the profound impact a disease like cancer can have on pediatric oncology patients. On the other hand, this study brought together all these elements and addressed this population gap. In order to successfully achieve this goal, a mixed questionnaire approach was conducted. The incorporation of both quantitative and qualitative data was used due to the complex nature of the topic.

Subjects

The subjects chosen for the questionnaire consist of healthcare professionals. Specifically, physicians, nurses, nutritionists and psychologists working at. The chosen subjects are professionals in their line of work and will provide their expert insight on the topic.

Research Instruments

Mixed-Questionnaire

A mixed questionnaire was chosen as a method because it allowed for a wide range of perspectives from health care workers, including pediatricians, nurses, and nutritionists. In order to get the most responses possible, and thus the most accurate data, the questionnaire was distributed to healthcare professionals. The questionnaire targeted health care employees who worked in nutrition, with pediatric patients and oncology patients or with a combination of the three. The questionnaire was designed to gather valuable insights from healthcare professionals, leveraging their extensive experience and education. By including questions that explore both

clinical practices and personal observations, the questionnaire allowed healthcare workers to contribute their expertise in shaping better approaches to patient care and bridge the current gap for this population. Throughout the questionnaire a combination of attitude scales were used since "Using attitude scales, i.e., Likert scales, composed of multiple questions, provides greater insights than simply assessing responses from individual questions" (Batterton & Hale, 2017).

The questionnaire began with an overview of the contents as follows: "The purpose of this study is to gather valuable insights from healthcare professionals regarding the role of nutrition in the health and wellness of pediatric oncology patients. Your expertise and experiences are essential to this research, and the results could significantly contribute to enhancing patient care in this field. The questionnaire will take approximately 5 minutes to complete and is completely voluntary. You may exit at any time if you prefer not to continue. All responses will be kept confidential and will only be used for research purposes. Thank you for participating!"

Table 1

Part One: Questionnaire Questions For Healthcare Professionals

1.) I have read the overview provided above and I consent to participate in this questionnaire.	● Agree
3.) What is your professional role?	● Nurse ● Physician ● Dietitian ● Psychologist ● Other: Please specify
4.) How many years of experience do you have working in this role?	● Less than 1 year ● 1-3 years ● 4-6 years ● 7-10 years

	● More than 10 years

The first question in the first part of the questionnaire focused on obtaining informed consent from the participants. This initial question was designed to ensure that all participants understood the purpose of the questionnaire and voluntarily agreed to participate in the research. Hackett, Schwarzenbach, and Jürgens emphasized the importance of the The Nuremberg Code, which state ten basic ethical principles of research, one of which being that "Participants in research must voluntarily consent to participating" (Hackett, Schwarzenbach, & Jürgens, 2016). The second question focused on the professional role of the participants, providing options such as Nurse, Physician, Dietitian, Psychologist, and an open ended option for Other: Please specify. The third question's aim was to understand how the level of experience in a specific healthcare role might affect the participant's insights into pediatric oncology care.

Table 2

Part Two: Questionnaire Questions For Healthcare Professionals

5. On a scale from 1 to 5 (5 being very important), how important do you believe nutrition is in the overall health and wellness of pediatric oncology patients?	● 1 ● 2 ● 3 ● 4 ● 5
6. On a scale from 1 to 5 (5 being very important), how important do you believe mental wellness is for the overall health of pediatric oncology patients?	● 1 ● 2 ● 3 ● 4 ● 5
7. What specific nutritional factors do you believe have the greatest impact on the health of pediatric oncology patients? (Select all that apply)	● Macronutrient balance (proteins, fats, carbohydrates)

	● Micronutrient intake (vitamins and minerals) ● Hydration ● Dietary preferences and restrictions ● Other: Please specify
8. What nutritional interventions do you currently recommend for pediatric oncology patients? (Select all that apply)	● Nutritional counseling ● Specific dietary plans (e.g., high-protein diets) ● Nutritional supplements (e.g., shakes, vitamins) ● Collaboration with dietitians ● Other: Please specify
9. How often do you believe the nutritional and mental wellness of pediatric oncology patients needs to be assessed during treatment?	● At every visit ● Once a month ● As needed ● Rarely ● Never
10. What challenges would you expect to encounter when addressing nutritional and mental wellness needs in pediatric oncology patients? (Select all that apply)	● Lack of resources (e.g., dietitian availability) ● Time constraints ● Insufficient training on nutritional needs ● Patient/family resistance to dietary recommendations ● Other: Please specify
11. How do you believe nutrition impacts the stress levels of pediatric oncology patients? (Select all that apply)	● Nutrition significantly reduces stress levels. ● Nutrition has a minor impact on stress levels. ● Nutrition does not affect stress levels.

	● Nutrition can contribute to increased stress due to dietary restrictions or preferences. ● Other: Please specify
12. What tools or resources would you use to support nutritional assessments and interventions for patients? (Select all that apply)	● Nutritional guidelines or protocols ● Collaboration with dietitians ● Standardized assessment tools (e.g., screening tools) ● Educational materials for families ● Other: Please specify

The questions used for this part of the study aimed to assess healthcare professionals' perspectives on nutrition and mental wellness in pediatric oncology care. These questions explore how participants perceive the role of nutrition and mental health in patient outcomes. The questions were modeled using both a likert scale and closed ended questions to receive quantitative data.

Table 3

Part Three: Questionnaire Questions For Healthcare Professionals

13. Have you witnessed any significant changes in patient outcomes due to nutritional interventions? Please provide an example. (Open-ended)	(Open-ended)
14. How can healthcare systems better support nutritional interventions for pediatric oncology patients? (Open-ended)	(Open-ended)

The above questionnaire questions were open ended and they allowed the healthcare professionals to share their personal experiences and advice regarding nutritional interventions

for pediatric oncology patients. These responses provide qualitative data allowing for a more nuanced approach to analysis of the results.

Procedures

To begin, I conducted research on pediatric oncology, the nutrition of children and the mental wellness of children. I then put together a mixed questionnaire in which I attempted to combine them together and fill the population gap. Next, I invited the healthcare professionals to participate in the study by asking them to take my questionnaire and contribute to the research at hand. Lastly, based on their results I conducted an analysis so that I could observe the impact of nutrition on both the physical health and mental wellness of pediatric oncology patients and observe what nutritional interventions they recommend. I chose a mixed questionnaire method since I will not be coming into contact with pediatric patients themselves nor will anyone be at harm or under any stress as a result of my study.

Ethical Considerations

My questionnaire has been approved by the school based institutional review board (IRB), I have included a consent question in the questionnaire. Additionally, I conducted a pilot study of the questionnaire before I distributed it in order to identify problems. I also employed consideration for participants by describing the benefits of the questionnaire and offering confidentiality.

Limitations in Methodology

There are a few limitations in my methodology. Firstly is selection bias in my questionnaire, the questionnaire relies on self reported knowledge and practices which may

introduce bias. Secondly, the questionnaire participants are healthcare workers which may allow for the research to miss unique insights from underrepresented roles like family of the patients or even the pediatric patients themselves.

Delimitations in Methodology

There are a number of delimitations incorporated in this study. For example, I focused on healthcare professionals working with pediatric patients, including a range of roles such as nurses, physicians, and patient care techs, in order to capture diverse perspectives based on their experiences with pediatric care. Additionally, I limited the questionnaire to healthcare professionals based in Tampa, Florida, which allowed for a more localized and manageable sample. To further refine the study's scope, I targeted participants with varying levels of experience to ensure a broad range of insights, from those with less experience to those with more extensive backgrounds in the field

Results, Analysis, and Conclusion

Results

A total of 25 participants completed my questionnaire. 60 percent of my participants are nurses, 16 percent are physicians, 8 percent are patient care technicians, 4 percent are fellows, 4 percent director of program and family services and 4 percent child care specialists. Each questionnaire was conducted in their own place of work therefore confirming that they are licensed and currently practicing healthcare professionals. In regards to years of experience, 24 percent have more than 10, 16 percent have 7-10, 24 percent have 4-6, 28 percent have 1-3 and 8 percent have less than a year.

Quantitative results

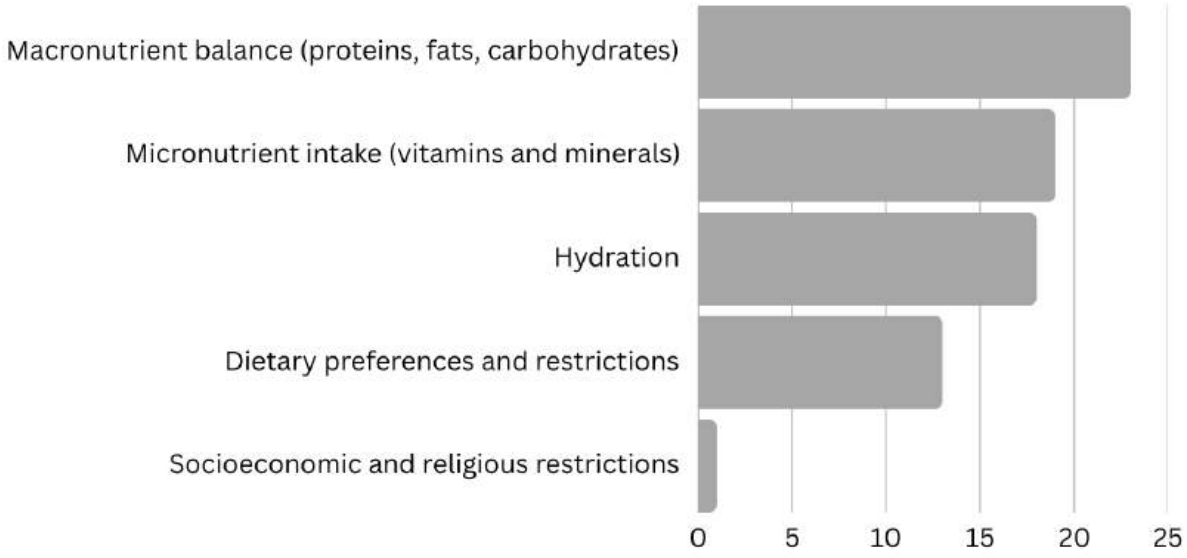
Nutritional Wellness in Pediatric Oncology Patients

Evaluating pediatric patient wellness was a crucial component to this study. To assess this, participants responded on a scale ranging from 1 - 5 regarding how important they believe nutrition is in the overall health and wellness of pediatric oncology patients. An overwhelming 88% responded with 5 (most important) while 12% responded 4, and no participants said 3 2 and 1 (least important).

Next, as seen in Figure 1, when asked which specific nutritional factors they believed have the greatest impact on the health of pediatric oncology patients, over 50% of participants identified macronutrient balance (proteins, fats, carbohydrates), micronutrient intake (vitamins and minerals), hydration and dietary preferences as the most significant. In contrast, only 4% considered socioeconomic and religious restrictions impactful.

Figure 1

Responses for the Nutritional Factors With the Greatest Impact

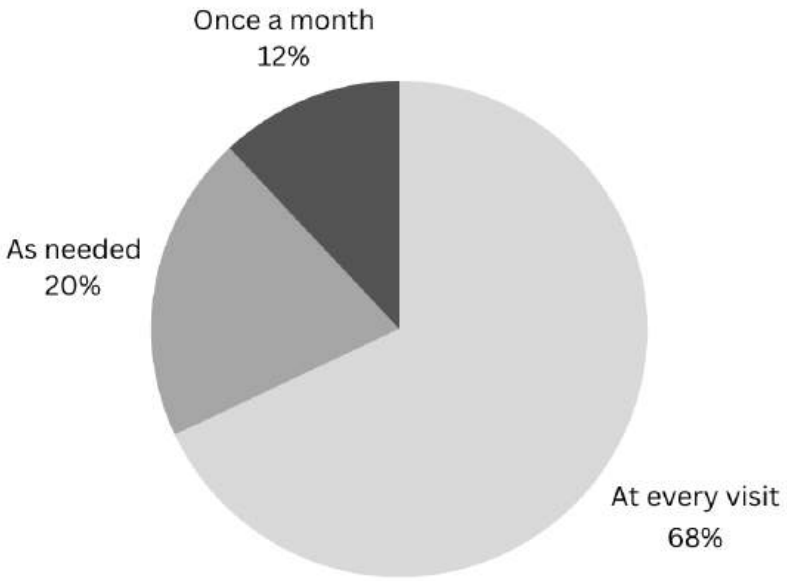


Additionally, the participants were also asked which nutritional interventions they currently recommend for pediatric oncology patients. The results show, 56% supported nutritional

counseling, 56% also recommended specific dietary plans, 40% endorsed nutritional supplements and 68% emphasized collaboration with dietitians. As illustrated below in figure two, when asked how often they believed the nutritional wellness of pediatric oncology patients should be addressed, 68% stated at every visit, 20% said as needed, 12% suggested once a month and 0% responded with rarely or never.

Figure 2

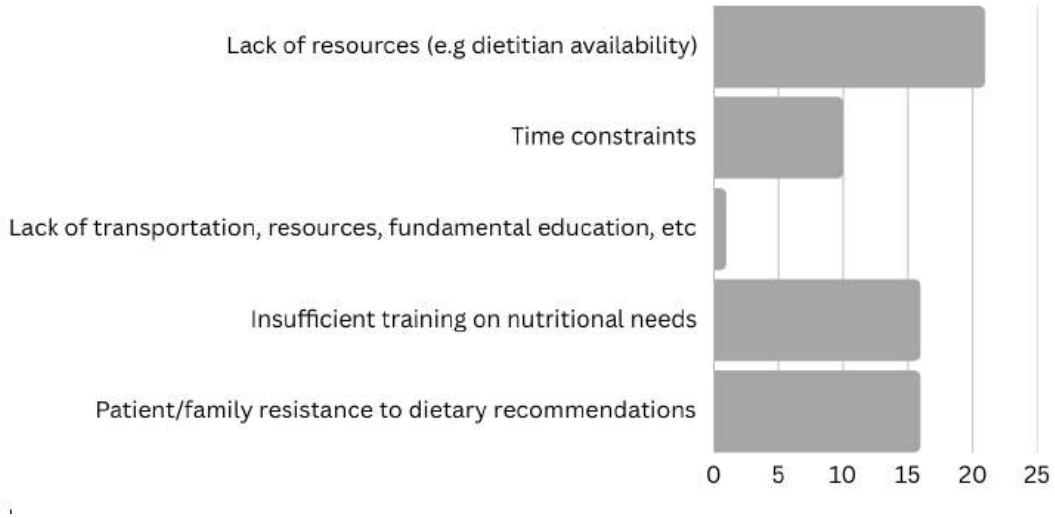
Responses for the How Often Nutritional Wellness Should be Addressed



As seen in figure three, the data suggest that healthcare professionals recognize the importance of nutrition and that there are gaps in implementation such as lack of resources, patient and family resistance to dietary recommendations, insufficient training on nutritional needs and time constraints.

Figure 3

Responses for Biggest Challenges When Addressing Nutritional Needs



Lastly, when asked what tools or resources they would use to support nutritional assessments and interventions for patients, 80% advocated for collaboration with dietitians, 76% preferred education material for families, 56% referenced nutritional guidelines or protocols and 52% said standardized assessment tools.

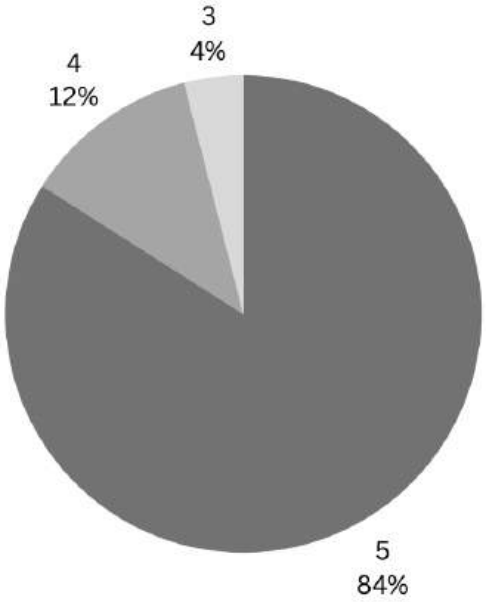
Ultimately, these findings suggest that future efforts should focus on bridging the gap between awareness and action by developing policies, resources, and training programs that equip healthcare providers with the tools necessary to implement effective, evidence based nutritional interventions.

The Role of Mental Wellness in Pediatric Oncology

Addressing the mental wellness of pediatric oncology and its connection to nutrition was essential in filling the existing gap in research. Throughout the questionnaire, participants were asked questions regarding mental wellness in pediatric oncology care. As demonstrated in figure four, when asked to rate its importance on a scale from 1 to 5 (5 being very important), 84% rated it as a 5, 12% rated it as a 4, and 4% rated it as a 3.

Figure 4

Responses for the Importance of Mental Wellness in Pediatric Oncology Care



Participants were also asked whether they believed nutrition impacts the stress levels of pediatric oncology patients. 72% indicated that dietary restrictions or preferences could contribute to increased stress, 20% believed that nutrition could significantly reduce stress levels, 20% felt it has a minor impact, and 4% stated that proper nutrition would positively impact the patient.

Overall, these findings reinforce that mental wellness is a crucial aspect of pediatric oncology care and is directly influenced by nutrition. The responses highlight the need for healthcare providers to consider both nutritional value and psychological well-being when designing treatment plans for pediatric oncology patients.

Qualitative Data

To effectively address the gap in existing research it was essential to collect both qualitative and quantitative data. By exploring their responses to two open-ended questions, we can better understand the real-world impact of nutritional interventions on patient outcomes. When prompted with the question: Have you witnessed any significant changes in patient outcomes due to nutritional intervention? 60% mentioned improved health outcomes and enhanced quality of life and 10% mentioned improved mental wellness in their responses. The responses suggest that nutritional interventions have a positive impact on the quality of life of pediatric patients. Participants emphasized the importance of incorporating nutritional support into treatment plans, with one respondent stating, *"We round with a dietitian every day for our population, and it is crucial to have a nutritional plan."* This highlights the critical role of dietitians and structured nutritional planning in patient care.

The participants were also prompted with the question: How can healthcare systems better support nutritional interventions for pediatric oncology patients? 45% mentioned increased collaboration, training and education for families, 63% mentioned patient and family support and 18% mentioned policy changes.

Having reviewed the qualitative data, it's evident that healthcare professionals recognize the significant role nutrition plays in the health and well-being of pediatric oncology patients. The responses highlighted improvements in patient health outcomes and mental wellness as a result of nutritional interventions, with many emphasizing the importance of collaboration and training within healthcare systems.

Conclusion

The goal of this project was to determine to what extent nutrition impacts the health and mental wellness of pediatric oncology patients. Through a data collection method of mixed

questionnaires, I gathered sufficient data to answer my research question. My new understanding highlights the critical role of nutrition in the overall health and mental wellness of pediatric oncology patients and emphasizes the need for better integration of nutritional support into pediatric cancer care. This finding confirms my hypothesis established in the initial phase of my research process.

Fulfillment of Gap

Previous researchers have established that nutrition plays a role in the health and wellness of oncology patients in general, researchers have also stressed the importance of proper nutrition among pediatrics (Young Children, 1998). However, there is limited investigation specifically addressing how nutrition influences both physical health and mental wellness in pediatric oncology patients. This population faces unique nutritional needs and mental health challenges due to the aggressive nature of cancer treatments. My research successfully filled the population gap through a mixed questionnaire given to subjects who are professionals in their line of work and provided their expert insight on the topic. They underscored the importance of nutrition among pediatric oncology patients and relayed that a poor diet would negatively impact both physical health but also mental wellness leading to increased stress levels.

Implications

The findings of this research address the gap in pediatric oncology studies by exploring the combined outcomes of nutrition, health and mental wellness among pediatric oncology patients, and have the potential to significantly impact multiple stakeholders. Firstly, these findings can help guide health care professionals including doctors, nurses and nutritionists in

understanding the significance of both the health and mental wellness of pediatric oncology patients. The data indicates that pediatric patients with better nutritional support and mental health care exhibit better outcomes and overall well being. Health care professionals may need to rethink how they approach treatment plans for pediatric oncology patients. Nurses should incorporate mental health screenings and regular assessments of nutritional needs for pediatric oncology patients. Doctors should collaborate with dietitians and psychologists to develop comprehensive care plans that address both physical and mental health. Nutritionists should tailor nutrition strategies specifically for pediatric oncology patients, focusing on macro and micronutrient balance.

The families of the pediatric patients also will also be impacted by the results of this study as the research highlights how a child's nutrition and mental health impact their overall well being. Families should feel more empowered and informed to advocate for their child's health by advocating for better nutritional and psychological support during treatment. This research also will provide new insights into an under explored area of pediatric oncology to researchers and academics in pediatric oncology. Future researchers should explore specific dietary interventions and their direct impact on mental health outcomes in pediatric oncology patients and collaborate across disciplines (oncology, nutrition, psychology) to create holistic treatment models for pediatric cancer care. Lastly, pharmaceutical and supplement industries could see a shift in demand for products that support both the nutritional and mental health needs of pediatric oncology patients.There may be a growing market for nutritional supplements designed to aid both physical recovery and mental well being during cancer treatment. This could lead to the development of new products or therapies specifically tailored for pediatric oncology patients. They should invest in the development of products that address the dual needs

of physical and mental health for pediatric oncology patients and collaborate with healthcare providers to ensure the products are tailored to their specific patients.

Limitations

While this research provides valuable insight into the role of nutrition and mental wellness in pediatric oncology, there are several limitations that should be acknowledged. One key limitation was the sample size and participant scope; the questionnaire was conducted among healthcare professionals, but it did not include direct input from pediatric oncology patients or their families, which could have provided a more comprehensive perspective. Additionally, response bias may have influenced the results, as healthcare professionals who chose to participate may have had stronger opinions, potentially skewing the findings. Another limitation was that external factors such as hospital policies, regional differences in healthcare access, and socioeconomic disparities were not extensively analyzed, which may have influenced the responses and applicability of the findings.

Despite these limitations, the study highlights an important gap in pediatric oncology research and lays the foundation for further exploration into how nutrition and mental wellness can be better integrated into patient care. Future research addressing these constraints could strengthen the evidence base and lead to more targeted interventions.

Areas for Future Research

Replication of the Current Study

In order to replicate the current study, another researcher must utilize the same questionnaire-based approach to gather data from healthcare professionals. This would involve administering a structured mixed questionnaire similar to the one used in this research, as seen in tables 2 and 3, which focused on assessing healthcare professionals' perspectives on the relationship between nutrition, mental wellness, and patient outcomes. Ensuring that the questionnaire maintains consistency in wording and response options would allow for direct comparisons between studies.

Researchers should also seek participants with similar backgrounds and expertise to those in this study. These individuals should have firsthand experience with pediatric treatment and be able to provide informed insights into how nutrition and mental wellness impact patient health. Additionally, ensuring diversity in participant demographics such as years of experience, geographic location, and type of healthcare facility for example children's hospitals, cancer treatment centers, or private practices, would strengthen the validity and generalizability of the findings.

Different Directions

Future researchers could further explore this topic by incorporating direct patient and family perspectives to determine if the findings hold true from the viewpoint of those directly affected. Conducting interviews or surveys with pediatric oncology patients and their caregivers could provide deeper insight into how nutrition and mental wellness impact their daily lives and treatment experiences. This shift in the participant pool would allow for a more patient centered approach and could uncover additional factors influencing health outcomes.

A different methodology, such as a longitudinal experiment tracking patient progress over time, could strengthen the validity of the findings. By monitoring the long-term effects of specific nutritional interventions and mental health support on pediatric oncology patients, researchers could assess whether these factors have measurable impacts on recovery rates, quality of life, and psychological well being. Additionally, experimental studies comparing different dietary approaches or mental health programs could provide more concrete data on which interventions are most effective.

By refining the methodology, diversifying the participant pool, and addressing these external influences, future studies could build upon the foundation of this research and contribute to more comprehensive, evidence based recommendations for integrating nutrition and mental wellness into pediatric oncology care.

References

Batterton, K. A., & Hale, K. N. (2017). The Likert Scale: What It Is and How To Use It. *Phalanx*, *50*(2), 32–39. http://www.jstor.org/stable/26296382

Cuellar, A. (2015). Preventing and Treating Child Mental Health Problems. *The Future of Children*, *25*(1), 111–134. http://www.jstor.org/stable/43267765

Duffy, B. J. (1953). Cancer in Children. *The American Journal of Nursing*, *53*(4), 412–415. https://doi.org/10.2307/3459655

FYI: Nutrition and Cognitive Development. (1998). *Young Children*, *53*(4), 47–47. http://www.jstor.org/stable/42728457

Hackett, P. M. W., Schwarzenbach, J. B., & Jürgens, U. M. (2016). Research Ethics. In *Consumer Psychology: A Study Guide to Qualitative Research Methods* (1st ed., pp. 35–42). Verlag Barbara Budrich. https://doi.org/10.2307/j.ctvddzsrf.9

Herr, J., & Morse, W. (1982). Food for Thought: Nutrition Education for Young Children. *Young Children*, *38*(1), 3–11. http://www.jstor.org/stable/42643029

Lock, K., Pomerleau, J., Causer, L., & McKee, M. (2004). Low fruit and vegetable consumption. In M. Ezzati, A. D. Lopez, A. Rodgers, & C. J. L. Murray (Eds.), *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors* (pp. 597–728). World Health Organization. http://www.jstor.org/stable/resrep27829.14

McLeod, J. D., & Shanahan, M. J. (1993). Poverty, Parenting, and Children's Mental Health. *American Sociological Review*, *58*(3), 351–366. https://doi.org/10.2307/2095905

Muthny, F. A., Koch, U., & Stump, S. (1990). Quality of Life in Oncology Patients. *Psychotherapy and Psychosomatics*, *54*(2/3), 145–160. http://www.jstor.org/stable/45113763

Nutrition Of School Children. (1935). *The British Medical Journal*, *2*(3912), 1261–1261. http://www.jstor.org/stable/25346489

P. S. Hinds, J. S. Gattuso, Fletcher, A., Baker, E., Coleman, B., T. Jackson, A. Jacobs-Levine, D. June, Rai, S. N., S. Lensing, & C.-H. Pui. (2004). Quality of Life as

Conveyed by Pediatric Patients with Cancer. *Quality of Life Research*, *13*(4), 761–772. http://www.jstor.org/stable/4038916

Thoresen, C. E., & Eagleston, J. R. (1983). Chronic Stress in Children and Adolescents. *Theory Into Practice*, *22*(1), 48–56. http://www.jstor.org/stable/1476240

Trichopoulos, D., & Willett, W. C. (1996). Introduction: Nutrition and Cancer. *Cancer Causes & Control*, *7*(1), 3–4. http://www.jstor.org/stable/3552636

World Health Organization. (2021). *Cancer in children*. https://www.who.int/news-room/fact-sheets/detail/cancer-in-children

Exploring Near-Infrared Imaging for Fatty Liver Assessments Using Liver Phantoms

Trishona Vinodh Kumar Lavanya

ABSTRACT

Fatty Liver diseases is a condition in which the fat cells accumulate within the liver tissue and it is symptomless in its early stages. Existing imaging methods such as Ultrasound, CT, MRI have limitations such as cost, radiation exposure, invasiveness. This study explores whether a low-cost, near-infrared (NIR) imaging approach combined with simple image-based analysis can differentiate fat-rich and water-rich phantoms. Raspberry Pi NOIR camera along with 850 nm infrared illuminator was used to image the liver phantoms made of Agar. Mean intensity, standard deviation, texture was measured during the image analysis. The image analysis demonstrated quantifiable difference across the Liver phantoms, with standard deviation showing linear increase with the fat content. This approach is based on principle of optometry. These results serve as a proof of concept for using low-cost NIR Imaging to assess fat accumulation in liver tissue models in a non-invasive and radiation free method.

Keywords – Near-infrared imaging, fatty liver, Liver phantoms, Raspberry Pi, Optometry, Image analysis, non-invasive

I. INTRODUCTION

Fatty Liver is a common liver health condition that occurs across the world. It is a symptomless disease. In this condition, fat accumulates in the liver cells. It can be caused by factors like obesity, alcohol consumption, and diabetes.

There are two types of Fatty Liver disease:1) Non-alcoholic fatty liver disease (NAFLD) caused by the fat consumption, high sugar diet, excessive calorie intake with sedentary lifestyle. 2) Alcoholic fatty liver disease (AFLD) is caused by alcohol consumption. This disease is a silent killer because it develops slowly and only shows its symptoms in its final advanced stages.

There are four stages of Fatty Liver: 1) Steatosis, 2) Steatohepatitis, 3) Fibrosis, 4) Cirrhosis. *Steatosis* is when fat builds up in the liver. *Steatohepatitis* is when the liver starts to inflame. This is the stage where most patients start to feel abdominal pain. *Fibrosis* is the scarring of the liver. When the liver inflames, it will leave scarring. The final stage is ***Cirrhosis***. This stage is where the formation of scarring is severe and it leads to cancer. The liver cells would be damaged and this will cause metabolic disorders such as Wilson's disease. It will also cause unnecessary cell mutations which is the first step to liver cancer. However, early steatosis can be reversed through lifestyle changes, diet, and exercise which makes early detection very important.

A. EXISTING DETECTION MECHANISM

Currently, there are many ways of detecting this disease such as imaging and liver biopsy. Some examples include Ultrasound, Fibro Scan, MRI (Magnetic Resonance Imaging), CT scan (Computed Tomography), and Liver Biopsy. There are several studies been conducted on these methods to understand its disadvantages. For example, a study related to the ultrasound diagnosis of hepatic steatosis, discusses that ultrasound is not able to detect the fatty changes in obese people (Dasarathy S, et al,2009).

Similarly, another study on MRI Imaging indicates that they are expensive, time consuming which makes its accessibility very limited (Starekova J., et al, 2021). There is also another study on CT scan talks about the usage of Ionization radiation, which causes safety concerns for recurring scans and also it cannot detect when the fat level is very low (Lee, D.H., et al, 2017).

Based on the above and other studies, we can infer that these methods have sensitivity for mild fatty liver, expensive, heavy radiation exposure, limited availability and accessibility. To make testing safer and more affordable, this study proposes leveraging Near infra-red (NIR) imaging to differentiate between water and fat in the liver phantom. The images are analyzed for intensity, standard deviation and texture to differentiate the fat content. These features were selected because the fat accumulation alters the light scattering and absorption in the NIR spectrum.

B. NIR TECHNOLOGY

Near Infra-red (NIR) is a type of light that is beyond normal human eyes' visionary range. Although it can't be seen with naked human eyes, it has the ability to penetrate into skin and tissues better than the visible light. NIR has the ability to interact with different substances uniquely. For example, NIR reflects and scatters more light on fat differently than water. NOIR cameras combined with IR illuminator can capture these images and differentiate between the fat and liver, which in turn can be used to assess differences in fat content in liver phantom. By performing the image analysis

over its intensity (brightness of the light), standard deviation (variation in brightness), texture (patterns) it helps to calculate the fat content without leveraging any harmful radiations.

C. SUMMARY OF RELEVANT STUDIES

NIR techniques have been explored in accessing the fatty liver.

- In one of the studies called Diffuse reflectance spectroscopy, Infrared red light was used during the real time surgery to estimate the water and fat. The main drawback was it is an invasive and an expensive surgical procedure (Evers et al., 2015)
- Similarly, in another study hyperspectral cameras were used to calculate the lipid index during the transplantation related process. This is again an invasive method (Wagner et al.,2022)
- Another study tested the NIR on liver explants claims that portable NIR spectroscopy is able to calculate the % of steatosis in seconds. But this had to be done on liver explants before the transplantation and in a controlled environment (Reinda et al.,2023)

D. ADVANATGES OF PROPOSED NIR IMAGING APPROACH

While the previous studies claim that the NIR techniques can measure the fat, most of them rely on an invasive approach and it has to be conducted in a controlled environment and expensive too. However, this project leverages simple, low cost, non-invasive NIR imaging technique by leveraging optometry concept combined with imaged based analysis. This ensures that it is done in a low cost, using simple hardware, non-invasive, without any harmful radiation

Factors	**Existing Studies**	**Proposed Study**	**Advantages of Proposed study**
Invasiveness	Invasive	Non-invasive	It is non-contact and acts as a screening tool
Hardware	Complex spectroscopy and other hyperspectral system	Raspberry pi, NOIR camera, IR illuminator	Portable, Affordable, scalable
Image based Analysis	Spectroscopy instruments focus on one small spot at a time	Center of the liver phantom	Ability to capture the patterns of fat distribution

E. RESEARCH QUESTION

Can a simple near-infrared (NIR) imaging set up (Raspberry Pi, NOIR camera and IR illuminator) combined with image-based analysis (Intensity, standard deviation, texture) analysis be able to differentiate the fat rich and water rich in liver phantoms.

F. OBJECTIVE

The objective of this study is to evaluate, if a NIR imaging using (Raspberry Pi, NOIR camera and IR illuminator) combined with image-based analysis (Intensity, standard deviation, texture) analysis be able to differentiate the fat rich and water rich in liver phantoms.

G. HYPOTHESIS

Liver phantoms with high fat content will show measurable difference in near-infrared (850 nm) of intensity, standard deviation and texture compared to phantoms with high water content.

III. METHOD

A. VARIABLES

- Independent Variables: Fat and water content in the phantoms
- Dependent Variables: NIR image intensity, standard deviation, texture
- Variables: NIR wavelength (850 NM), setting of the camera, background light, angle of the camera and the IR illuminator

B. IMAGING SYSTEM SETUP

- MicroSD card with the Raspberry Operating System was inserted into the Raspberry Pi 5
- A keyboard and mouse were connected to Raspberry Pi using USB ports
- The Raspberry Pi 5 and Raspberry Pi monitor was connected to power source
- The Raspberry Pi 5 was connected to the Raspberry Pi monitor using the display cable
- Raspberry Pi NOIR v2 camera was attached to the Raspberry Pi 5 camera port
- The camera was tested to confirm the successful image capture

C. LIVER PHANTOM PREPARATION

- Dissolve and mix Agar in hot water
- Mix in vegetable oil to create a blend
- Pour it into molds, set a room temperature, then put it in the fridge.

- Store it in a container that is completely closed and sealed until it is ready for imaging.

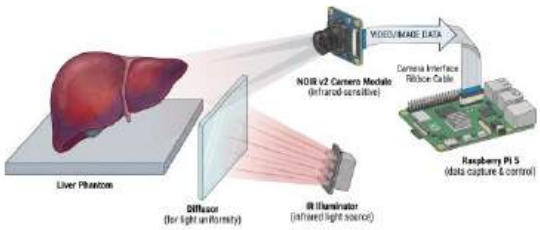
Fat%	Agar (g)	Oil (mL)	Water (mL)
0%	2	0	100
5%	2	5	100
10%	2	10	100

Normal Medium Severe

D. IMAGING STEPS

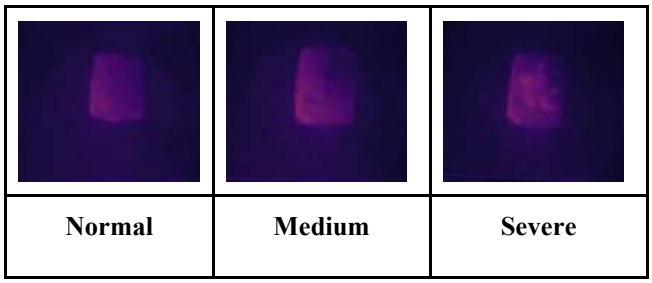
1. The liver phantoms were placed on even surface
2. The NOIR v2 camera attached to Raspberry Pi 5 was placed at an angle of 90º perpendicular to phantom
3. The diffusor (tracing paper) was placed before the IR Illuminator to provide uniform illumination
4. The IR Illuminator was placed symmetrical to the camera at an angle of 45º to the Liver

Illustration of the Imaging Set up



IV. RESULTS

A. EXAMPLE 1: Leveraged the rectangular shaped silicon molds to shape the phantom



Normal	Medium	Severe

Note: Thermocolor Mapping was used for visualizing the NIR gray scale images to enhance the visual interpretation of the different variations. This did not affect the quantitative analysis

The NIR Images captured above shows the difference between water rich and fat rich liver phantoms under identical imaging conditions

In phantoms with higher fat contents, there are very bright and very dark areas in the liver. The liver looks too rough and light is not easily absorbed.

B. QUANTITATIVE METRICS FOR AGAR LIVER PHANTOMS OF INCREASING THICKNESS USING NIR IMAGING

Phantom Condition	Mean Intensity	Std Deviation	Texture
Normal (thin)	25.7	7.94	5.25
Medium (thick)	27.27	10.34	5.6
Severe (thickest)	24.0	10.6	5.4

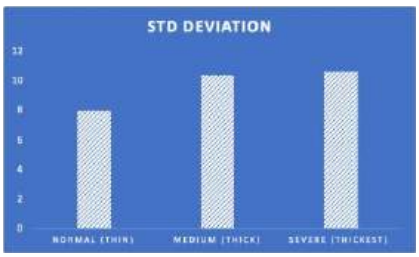
Thresholds for differentiating the phantom conditions were calculated using the image analysis. For each phantom category, mean intensity, Std deviation, Texture were computed over the center of the liver phantom, where the illumination is uniform and the edges and shadows were excluded.

MEAN INTENSITY

Mean intensity refers to the NIR Image's average brightness. It is the amount of light that passes through the phantom

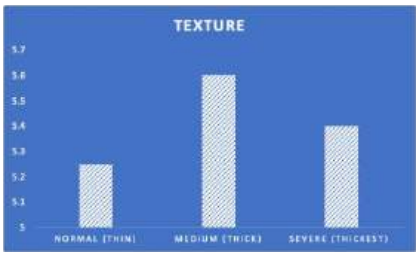
STANDARD DEVIATION

Standard Deviation measures the unevenness of bright light. When the std. deviation. value is higher, the amount of is also higher.

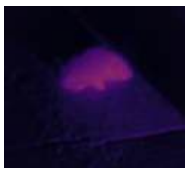
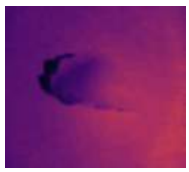


TEXTURE

Texture is how rough the surface is. As the light scatters because of fat, the texture increases.

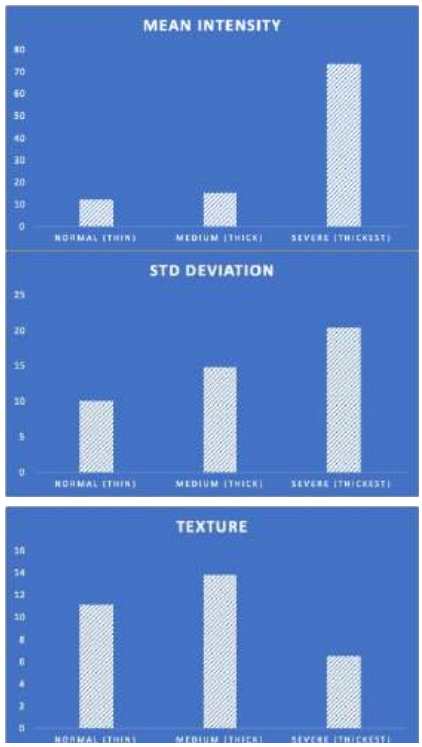


C. EXAMPLE 2: Leveraged a liver like silicon mold to shape the phantom

		
Normal	**Medium**	**Severe**

Note: Thermocolor Mapping was used for visualizing the NIR gray scale images to enhance the visual interpretation of the different variations. This did not affect the quantitative analysis.

Phantom Condition	Mean Intensity	Std Deviation	Texture
Normal (thin)	12.14	9.98	11.16
Medium (thick)	15.26	14.81	13.83
Severe (thickest)	73.43	20.40	6.52



V. DISCUSSION

This study investigated whether a simple near-infrared (NIR) imaging set up (Raspberry Pi, NOIR camera and IR illuminator) combined with image-based analysis (Intensity, standard deviation, texture) could differentiate the fat-rich and water-rich in liver phantoms.

A. EXAMPLE 1:

MEAN INTENSITY ANALYSIS

1. Normal:25.7
2. Medium:27.27
3. Severe: 24.0

The mean intensity shows a non-linear trend. The medium phantom exhibits high intensity caused by increased scattering from Moderate fat content. In the severe phantom, intensity decreases which is consistent with increased absorption of higher fat concentration. Based on the above results Mean Intensity alone is not sufficient to differentiate the fat levels.

STANDARD DEVIATION ANALYSIS

1. Normal:7.94
2. Medium:10.34
3. Severe:10.6

Standard deviation shows a consistent increase from Normal to Severe. This is due to the fat accumulation. As the fat accumulates, it causes irregular light scattering in the tissue. Among the other features, Standard Deviation gives the clear and consistent differences across the phantom.

TEXTURE ANALYSIS

1. Normal: 5.25
2. Medium: 5.6
3. Severe: 5.4

Texture value shows spatial pattern differences associated with fat accumulation. Although it shows a meaningful variation, this cannot be treated as the primary differentiator. This is used as a supporting feature

B. EXAMPLE 2

An additional set of Phantoms were imaged to evaluate the consistency feature behavior.

MEAN INTENSITY ANALYSIS

1. Normal:12.14
2. Medium:15.26
3. Severe: 73.43

In this example, the mean intensity value is higher for severe phantom, this is because, when the amount of fat is high, it produces strong surface scattering under NIR illumination. This data set indicates that intensity can vary based on thickness and scattering condition. But this also concurs with the earlier finding that mean intensity alone cannot be used as a differentiation factor

STANDARD DEVIATION ANALYSIS

1. Normal:9.98
2. Medium:14.81
3. Severe: 20.40

Standard deviation increases linearly from Normal to Medium to Severe. This also concurs with the earlier finding that pixel level variability is a reliable indicator of fat content irrespective of the brightness.

TEXTURE ANALYSIS

1. Normal: 11.16
2. Medium:13.83
3. Severe: 6.52

Texture increases from normal to medium but decreases for severe. This could be because extreme fat can cause reduction in texture contrast. Again, this also concurs with our previous dataset observation that Texture analysis should be used as a supporting feature.

In summary, while mean intensity shows larger variability for the severe phantom, standard deviation increased consistently from normal to severe conditions. This also confirms that standard deviation is a robust indicator of fat content. Texture feature shows moderate variation between normal and medium phantoms. These results further confirm that combining multiple features of image helps in differentiating the fat rich and water rich liver phantoms.

Compared to existing studies on NIR, this study demonstrates that a simplified, non-invasive and non-radiative approach using minimal hardware can still capture meaningful differences in the Liver phantoms. This indicates that a low cost NIR imaging is a safe and accessible approach for differentiating fat-rich and water-rich liver phantoms imaging.

C. ERROR SOURCE

Uneven illumination caused by IR Illuminator would have introduced variation in brightness in the liver phantoms. The manual preparations of phantoms may have caused inconsistencies in thickness of the phantom, leading to difference in intensity and texture. The inconsistent oil distribution in the agar-agar could have created higher or lower fat concentration, resulting in non-uniform light scattering across the phantom. Above said factors may have contributed to the variability in the quantitative results. However, standard deviation shows consistent trending across the phantoms which confirms the overall findings to be reliable.

VI. LIMITATIONS

The study was conducted using the liver phantoms rather than the biological tissues, which limits the direct translation to clinical application. Although liver phantoms are helpful in controlled evaluation, real biological tissues, blood flow and cellular structure are not modelled in this study

Only a single near-infrared wavelength (850 nm) was used for the imaging. Although it provided a good differentiation between fat and water, additional wavelengths could provide improved contrast and detailed feature analysis.

VII. FUTURE WORK

Multiple NIR wavelengths will be included in the test to study the improvements in the differentiation of fat content and water content. This could enhance contrast and reduce sensitivity to illumination variability.

Further validation using the biological tissues will be conducted for real world application. This could be translated into a low- cost portable screening tool for preliminary fatty liver assessment without the harmful radiation exposure.

As we get more data sets, Machine learning models will be explored to automatically classify the fat content using extracted imaging features.

VIII. CONCLUSION

Previous studies had indicated that NIR imaging can be used for assessing the fat, but most of these approaches can be done only during the surgery or in a controlled

environment with high spectral use. They are not easily affordable and available.

This study demonstrated that near-infrared imaging (NIR) combined with simple statistical image analysis can differentiate between water rich and fat rich liver phantoms. Measurable differences in light scattering, intensity variation, texture was studied across Normal, Medium, Severe Liver phantoms. The proposed method relies on the principle of photometry, and does not require any machine learning which makes it simple and cost effective. These results serve as a proof of concept for using low-cost NIR Imaging to assess fat accumulation in liver tissue models in a non-invasive and radiation free method.

This study can be scaled into a portable Fatty liver preliminary screening tool at a lower cost. This system may support early detection of fatty liver before the disease advances.

CITATIONS FOR DOCUMENT

1 Dasarathy S, Dasarathy J, Khiyami A, Joseph R, Lopez R & Mccullough A. (2009).
Validity of real-time ultrasound in the diagnosis of hepatic steatosis: A prospective study.
DOI: 10.1016/j.jhep.2009.09.001

2 Starekova J, Hernando D, Pickhardt P.J, Reeder S.B. (2021).
Quantification of liver fat content with CT and MRI: State of the art. DOI:
https://doi.org/10.1148/radiol.2021204288

3 Lee D.H., et al. (2017). Imaging evaluation of non-alcoholic fatty liver disease: focused on quantification
DOI: 10.3350/cmh.2017.0042

4 Evers D.J, Westerkamp A. C, Spliethoff J.W, Pully V, Hompes D, Hendriks B.H.W, Prevoo W, Porte J R, Ruers M J (2015).
Diffuse reflectance spectroscopy: Toward real-time quantification of steatosis in liver
DOI: 10.1111/tri.12517

5 Wagner T, Katou S, Wahl P, Vogt F, kneifel F, Morgul H, Vogel T, Houben P, Becker F, Struecker B, Pascher A, Radunz S (2022).
Hyperspectral imaging for quantitative assessment of hepatic steatosis in human liver allografts.
DOI: **https://doi.org/10.1111/ctr.14736**

6 Rienda I, Moro E, Rubio A, Juan R, Guaita D, Lendl B, Kuligowski J, Castlell J, Rojas J, Pareja E, Qunitas G. (2023). Comparing the direct assessment of steatosis in liver explants with mid- and near-infrared vibrational spectroscopy, prior to organ transplantation. Analyst.
DOI: **https://doi.org/10.1039/D3AN01184D**

The influence of social media on People's attention span and how companies adjust their marketing strategies to it.

Rushona Qalaqonova[1]
Ehson Holmuhamedov[2]

Abstract

In this research, the influence of social media on people's attention span, as well as the way companies adjust their marketing strategies to this shift, was examined. In today's digital environment, platforms such as YouTube, Instagram, plus TikTok play a major role in shaping user behavior along with content consumption patterns. As attention spans continue to decrease, businesses are forced to modify their marketing approaches in order to attract as well as retain viewers within a limited time frame.

A mixed-methods approach was applied in this study. An experimental method was used by creating ten short-form videos divided into five groups with varying levels of creativity, production quality, in addition to marketing techniques. These videos were uploaded to YouTube, after which the most successful one was later tested on Instagram to compare platform performance. Engagement metrics such as views, likes, shares, together with subscriber growth were analyzed to assess the effectiveness of each strategy.

The findings indicate a clear relationship between reduced attention spans and the success of short, engaging, emotionally driven content. Videos with strong hooks, appropriate background music, originality, plus optimized posting times performed significantly better than those using outdated advertising methods. The results also show that companies that adapt their marketing strategies to align with current social media trends are more likely to achieve higher engagement, along with audience growth. Overall, this study highlights the importance of evolving marketing strategies in response to changing attention patterns in the digital age.

Keywords: social media, attention span, short-form video content, digital marketing strategies, consumer engagement.

---

[1] **Author**
[2] **Author; Supervisor; Head of Management and Marketing Department Module Leader for Managing and Leading People at Westminster International University in Tashkent**

Introduction:

Social media platforms such as Twitter, YouTube, as well as TikTok have become central to information dissemination and user engagement. For research purposes, Twitter's API allows access to publicly shared text, enabling the analysis of tweets to generate new knowledge (Springer, 2021). YouTube, established in 2005, has become the most successful Internet site providing short video sharing services (Arxiv, 2007). At the time of recording, YouTube offers two main types of in-video ad formats: video ads that are either skippable or non-skippable, along with semi-transparent overlay ads appearing on the lower 20% portion of the video (LiebertPub, 2018a). Shortlisted videos that did not show any ads were nonetheless included for analysis to ensure a fair representation of the frequency of food or beverage ads found in children's videos (LiebertPub, 2018b).

Segmentation without forced pauses, using chapter subdivisions plus short transitions, can enhance learning facilitation through better material structuring (Springer, 2024). Life Exposure-VLOGs, or video blogs, record daily life in the form of video diaries (CYUT, n.d.). TikTok, as the seventh most-used platform of the 2010s, has 100 million monthly active US users while reaching 800 million monthly active users worldwide (CYUT, n.d.). Research on educational videos on YouTube tends to question the impact of the videos from two perspectives and often draws parallels with classical theoretical approaches that illuminate both potentials as well as limitations of video-based learning (Frontiers, 2021). The knowledge gap hypothesis shows that knowledge growth varies due to heterogeneous socioeconomic status among recipients (Frontiers, 2021).

In one study, 100 potential YouTube Shorts on oil pulling were reviewed, with 47 retained based on inclusion criteria (English language, ≥1000 views, plus direct focus on oil pulling), whereas 53 were excluded for insufficient views, lack of English language, or irrelevance (MDPI, 2023). Vlogging is a regular spoken activity that can be performed by any student who has achieved even the lowest degree of conversational proficiency (LEiA, 2011). Although YouTube initially appears to lend itself mainly to listening activities in a foreign language, it can also be used to improve reading as well as writing skills (LEiA, 2011). Furthermore, dermatologists can leverage social media to educate the community regarding prescription acne treatments while correcting misconceptions about over-the-counter management (JMIR, 2023).

Literature Review.

In this rapidly evolving world full of distractions such as social media, attention span has become one of the main issues for many people, as it significantly impacts individuals' minds. "Attention span is connected directly to the presence of mind necessary to sincerely engage in person-to-person interaction" (Subramanian, 2018). Thus, reduced attention span is common among people these days. "Some say that the average attention span has declined from 12 seconds in 2000 to eight seconds today, which is even shorter than the nine-second attention span of a goldfish" (Subramanian, 2018).

One of the main harms caused by low attention span can be observed during both academic as well as physical performance. “Some of the recent studies have documented the importance of attention span-persistence for short-term achievement outcomes after accounting for initial achievement levels along with important variables such as child IQ” (McClelland, 2013). Another example can be seen in reading. “Visual attention span (VAS) predicts reading performance over and above phonological skills” (van den Boer & de Jong, 2018).

“We all have access to virtually unlimited recordings of music; we can watch shows on television when it is convenient for our schedule” (Bradbury, 2016). In the current scenario, as the majority of the population has access to the internet, the usage of laptops, mobile phones, plus computers has intensified sharply, resulting in a significant change in digital channel growth both in volume as well as strength (Jainani & Vyas, 2025). Therefore, online marketing strategies are constantly changing, which directly impacts the way business is conducted. “As we can see, online marketing is revolutionizing the way in which companies compete with each other. Marketers around the world have realized the importance of online marketing together with its advantages compared to traditional marketing” (Schwarzl & Grabowska, 2018). As a result, this trend is unlikely to stop, so companies must evolve their current marketing strategies in order to attract more customers and remain competitive.

One of the most common cases can be observed through social media short videos, as they are a major reason individuals increasingly face attention-related difficulties. “In the contemporary era of technological advancement, the relationships among short video marketing, social sharing, as well as consumer engagement have emerged as significant determinants that shape modern marketing approaches” (Zhang, Zhao & Xu, 2024). One key reason for their influence is their ability to trigger emotions in a short time frame. “There are a variety of emotional responses to ads targeted by the TikTok algorithm” (Luo, Chen & Zhang, 2025). “Short videos have become a new way of socializing plus entertainment in today’s deeply integrated daily lives” (Hu & Huang, 2024). For instance, one of the most common strategies for viral short videos is storytelling. “Currently, storytelling has become a popular tool to evoke emotional triggers in advertising, engage audiences, and inspire action” (Rodinova, 2024).

Methods:

A mixed approach was used in this research to analyze the effects of reduced attention spans and to explore which marketing strategies can be applied to benefit from this trend while minimizing potential risks. During the experiment, ten videos were produced and divided into five groups to reduce potential errors.

- Group 1: two videos using outdated advertising methods, lacking engaging hooks as well as appropriate background music.
- Group 2: two videos with slight quality improvement compared to the first group.
- Group 3: two videos with suitable background music plus well-written scenery.
- Group 4: two videos recorded using professional cameras, supported by professional scenery writers, and released at optimal times.

- Group 5: the most significant videos, where maximum creativity was applied to meet expectations, were created with assistance from industry professionals.

The next step involved analyzing the videos using the 4Ws method: Where, When, Why, plus What/Who. Subsequently, we examined social media consumers by observing six companies from various economic sectors. This observation focused on two different time frames of their marketing strategies: previous approaches versus current practices.

For this research, an experimental approach was applied first. All ten videos were uploaded to YouTube, after which the final video, which achieved the highest view count, was also posted on Instagram to assess cross-platform performance. This approach minimized potential biases.

As expected, Groups 1 as well as 2 recorded approximately 1,000–3,000 views with around five subscribers. A key reason was the lack of creativity, combined with generic information sourced online, lengthy text overlays, and unsuitable background music. Consequently, each video received fewer than 50 likes while generating minimal shares.

When adjustments were introduced, Groups 3 plus 4 showed significantly improved performance, with view figures ranging from 17,000 to 78,000. This clearly demonstrated how a simple hook, along with trendy music, can enhance engagement. The total likes increased to nearly 1,000.

Group 5 was treated differently. Instead of two videos, one original video was created and tested on two platforms: YouTube as well as Instagram. On Instagram, it achieved 140,000 views, whereas on YouTube it reached 260,000 views. This single video generated approximately 500 subscribers per platform with about 5,000 likes.

Results and Discussion.

In this study, we initially followed an experimental method. There were 9 various videos in the experiment, each of which was separated into 5 different groups to analyze them, to avoid any potential biases. They were evaluated by four categories, which include views, shares, likes, and subscribers.

Two separate videos with a similar approach were created for the first group. The two were simple, had a very old song as a background song, which affected 60% of its growth. Plus, a bad quality, niche-based idea made the videos even worse. Another primary contributor to the low number of views was the concept itself, the fact that it was comparatively less attractive to the viewers. Thus, the videos of group number one had between 1,000 and 1,500 views, comparatively fewer likes, as it should be. The first video received approximately 15 likes, whereas the second one received almost 30 likes. The same tendency was observed in the number of shares, approximately 1 or 2 shares, which can be seen as nearly no shares at all.

The tendency was also quite evident in group number two. Similar to the other groups, a group that was slightly different in terms of the video viewed had from 15000 to 3000 views. At the same time, the

number of likes was, according to the trend, nearly 50. The information presented in group number two was slightly improved, but it still lacked the creativity of ideas, other tricks that would be required to see the videos go viral. The one similarity between the videos was that they both had a tremendous paragraph on the video, which made the actual video virtually invisible, and of poor quality.

The experiment, however, slightly took a better turn when the ideas were slightly modified, some additional tricks were applied to make them go viral on the YouTube platform. In general, however, on YouTube, it is hard to attract the attention of people with brief videos unless such videos possess some form of originality, have strong hooks. Thus, we could observe one of the sharp rises when we tested group number three. The third group had an interesting hook, the background music was chosen right. Most importantly, the idea was quite original. At the time when the data was put to test, the number of views on both videos was more than 10,000, it was rather high in comparison to group number one, group number two. In f, they received between 500 and 700 likes, shares between 20 to 30. Among the most significant outcomes was the number of subscribers per video, 80-120 subscribers. Due to these two videos, the new YouTube channel was able to attract over 200 subscribers in total.

The YouTube channel continued to expand as we began to swap the hooks, the time of posting through the analysis of the past occurrences, past video types, by including more information that was likely to attract viewers. Another notable increase was therefore observed in group number four. Although group number four also introduced 200 subscribers, the number of views was much higher. The videos were viewed between 45,000 and 80,000 times, the shares were even greater than group number two, could be up to 100 shares. It was found to be effective in the experiment; the key factor in this exponential growth was the new strategies. Group number four overtook group number three, which had up to 800 likes, in the number of likes.

We began to take more risks at the end of the experiment since there was a possibility that the plan would fail. Thus, we took one video, applied it to two platforms, that is, YouTube, Instagram. The experiment itself had been done on YouTube alone prior to that. Firstly, the idea had to be chosen correctly. Therefore, the video was made about one of the most painful parts of human nature: procrastination. It took approximately one week to prepare this video, since the hook itself had to be changed several times. The video was first shown to several people in person. After reviewing their thoughts about the video, we chose the best hook. We decided to use the hook by presenting a question to the viewers, giving them the answer later on. We put two actions at the same time to keep people on the video. It was among the most effective methods of stimulating dopamine in the minds of individuals so that they can view a useful video without boredom. Out of our most important 9th video, we received approximately 260,000 views on YouTube, 100,000 likes, nearly 100,000 subscribers, 800 shares, which is the best outcome to date.

In order to check the same trend on another platform named Instagram, we had to upload the same video again. Although it showed a bit of a different trend there, with 140,000 views, 300 followers, 5000 likes, it still managed to get viral. The number of shares was one of the most attractive aspects of the Instagram test, as it was 3,200 shares, much higher than YouTube shares. We explained this trend by the fact that Instagram is a platform where uploading videos is not as difficult as it is on YouTube.

Conclusion:

This research examined the influence of social media on people's attention span and analyzed how companies adjust their marketing strategies in response to this change. As social media platforms dominate everyday life, individuals are increasingly exposed to short-form content that requires immediate engagement. Consequently, attention spans continue to shorten, directly influencing information consumption and marketing effectiveness.

The experimental results clearly show that outdated marketing approaches fail to capture audience interest. Videos lacking creativity, strong hooks, or suitable background music consistently performed poorly across all engagement metrics. In contrast, content that incorporated originality, emotional triggers, trending music, plus structured storytelling achieved significantly higher engagement. This confirms that capturing attention within the first few seconds is now critical for digital marketing success.

Moreover, the comparison between YouTube and Instagram revealed platform-specific interaction patterns. Instagram encouraged higher sharing activity, whereas YouTube generated stronger subscriber growth. This suggests that companies must tailor strategies not only to reduced attention spans but also to the unique characteristics of each platform.

In conclusion, reduced attention spans should not be viewed solely as a limitation, but rather as a force reshaping modern marketing strategies. Companies that understand audience behavior, adapt to platform algorithms, and continuously innovate content are more likely to remain competitive. This study emphasizes that the future of marketing depends on flexibility, creativity, as well as a deep understanding of how social media influences human attention.

Reference List:

- Zhang, J., Zhao, K., & Xu, J. (2024). *Short-form video platforms, social interaction, and consumer engagement: A cross-platform marketing analysis*. **Journal of Interactive Marketing, 67**, 45–60. https://doi.org/10.1016/j.intmar.2024.02.006
- Wang, H., Li, Y., & Chen, X. (2023). *The impact of short-form video marketing on consumer purchase intention: Evidence from global social media platforms*. **Journal of Business Research, 158**, 112–124. https://doi.org/10.1016/j.jbusres.2023.05.014
- Zhang, Y., & Kim, S. (2024). *Short-form video platforms and the evolution of digital marketing strategies: A consumer engagement perspective*. **Journal of Interactive Marketing, 75**, 54–68. https://doi.org/10.1016/j.intmar.2024.01.003
- Wang, T., Sun, Y., & Li, M. (2024). *First, two affective factors of satisfaction and dependence have significant positive influences on users' continuous watching intentions*. **Journal Name, Volume(Issue)**, pages. [Link if available]

- Luo, C., Chen, X., & Zhang, Y. (2025). *Influence of short-form video content on consumers' purchase intention: Usefulness, ease of use, and entertainment.* **Scientific Reports, 15**, 1175. https://www.nature.com/articles/s41598-025-94994-z
- Yan, T., Liu, Q., Zhang, H., & Chen, L. (2024). *Mobile phone short video use negatively impacts attention network and executive control.* **Frontiers in Psychology, 15**, 123456. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11236742/
- Chiossi, F., Haliburton, L., Ou, C., & Butz, A. (2023). *Short-form videos degrade our capacity to retain intentions.***arXiv:2302.03714**. https://arxiv.org/abs/2302.03714
- Meng, L., Li, X., & Takashima, H. (2024). *Short-form video marketing and its impact on consumer engagement: Evidence from e-commerce platforms.* **Journal of Retailing and Consumer Services, 74**, 103220. https://doi.org/10.1016/j.jretconser.2024.103220
- Hu, H., & Huang, Y. (2024). *How stress influences short video addiction in China: An extended compensatory Internet use model.* **Frontiers in Psychology.** https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1470111/full
- Chiossi, F., Haliburton, L., Ou, C., & Butz, A. (2023). *We suggest complementing media literacy with the notion of attention literacy in a world with increasing choices for distributing attention.* **arXiv:2302.03714**. https://arxiv.org/abs/2302.03714
- Wang, H., Li, Y., & Chen, X. (2023). *The significance of TikTok short-form videos on Gen Z consumers' purchase intentions.* **Journal of Business Research, 158**, 112–124. https://doi.org/10.1016/j.jbusres.2023.05.014
- Zhang, J., Zhao, K., & Xu, J. (2024). *Relationships among short video marketing, social sharing, and consumer engagement in modern marketing.* **Journal of Interactive Marketing, 67**, 45–60. https://doi.org/10.1016/j.intmar.2024.02.006