thurj

the harvard undergraduate research journal

Fall 2025
Vol. 16, Issue 1

thurj
the harvard undergraduate research journal

Fall 2025 Best Manuscript

Adnan Bin Alamgir
Golam Yasin

Adnan Bin Alamgir '29 is a first-year at Harvard College. Originally from Bangladesh, he is a prospective Physics major with a secondary in Arts and Visual Studies. His academic interests lie in the interdisciplinary connection between physics, environment, and the arts. Outside of academics, Adnan is a member of the Intellectual Vital Student Board and active in theater and fencing. In his free time, he loves playing the piano and cycling.

Golam Yasin is a high-school gap-year student from Bangladesh who graduated from Dhaka Residential Model College. His academic interests span biology and environmental science. Through his contribution to "Otute" (a startup designing sustainable bricks) and his ongoing research on water treatment and disaster resilience, he is committed to applying scientific insight to real-world needs. In his free time, he loves painting and game development.

Submit your research to

thurj the harvard undergraduate research journal

THURJ is the only campus publication that showcases peer-reviewed undergraduate student research.

Circulation to dorms, departments, libraries, and scientific institutions.

Submit your research, manuscript or thesis excerpt!

Reference www.thurj.com/submit for submission guidelines.

THURJ accepts submissions that reflect original research in all disciplines.

$200 for Top Manuscript!

For more info, including our latest issues, visit www.thurj.com.

To contact the Executive Board, contact thurjall@gmail.com or

visit https://www.thurj.com/executive-board.

December 2025

Dear Harvard Community,

We are delighted to present the Fall 2025 issue of *The Harvard Undergraduate Research Journal* (THURJ), a student-run, biannual publication committed to showcasing and celebrating exceptional research from Harvard undergraduates of all disciplines and interests.

Founded in 2007, THURJ established itself as a hub for undergraduate research until the COVID-19 shutdown. Revived in Fall 2023 under Editor-in-Chief Ellie Shin '25, it released its first post-pandemic print edition. Two years later, THURJ is more active than ever and among Harvard's largest undergraduate organizations.

In a time of growing polarization and distrust toward science and academia, we have set out to play a small part in strengthening scientific literacy and highlighting the importance of research and discovery, even in all its complexities. Our theme, *Mapping Knowledge and Complexity*, reflects this commitment. To support this effort, we launched a new position—Director of Education and Engagement, held by Héctor Martínez Luna '27, who led three workshops for our team—and we've invited Dr. Insoo Hyun, Director of the Center for Life Science and Public Learning at the Museum of Science, to present on science communication to the THURJ team.

We extend our gratitude to every member of THURJ for their dedication and commitment, and to our faculty reviewers, whose insights are essential to maintaining the quality of our publications. We also thank our faculty advisor, Dr. Andrew Berry, who embodies our belief in the importance of science communication and research engagement for the next generation of scientists and leaders. Finally, we are grateful to the Department of Science Education, Harvard Medical School, the School of Engineering and Applied Sciences, the Department of Arts and Humanities, and many others for their invaluable support.

In this edition, we feature a broad range of fields from cancer biology and maternal health to novel filtration technologies as well as congressional racial dynamics, with five manuscripts of original research by Harvard undergraduates and seven insightful commentaries on the cutting edge of research at Harvard and beyond. We hope that you enjoy reading and learning from these excellent pieces as much as we have.

Sincerely,

Aditya Tummala
Editors-in-Chief

Theo Tobel

Table of Contents

Research

Features

Licensing

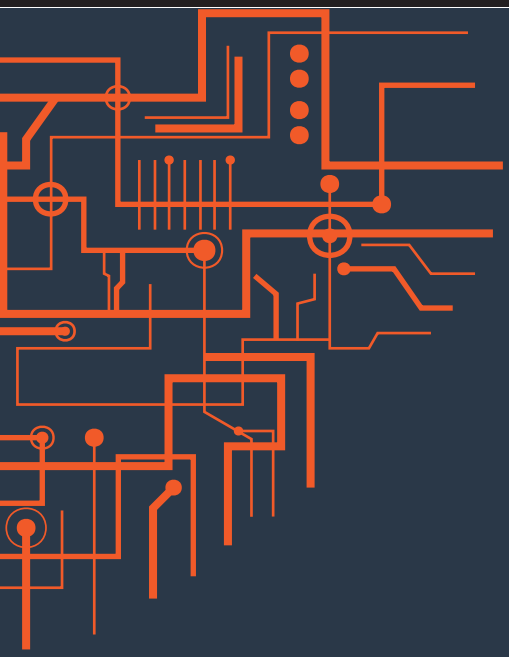
About Us

The Harvard Undergraduate Research Journal (THURJ) showcases peer-reviewed undergraduate student research from all academic disciplines. As a biannual publication, THURJ familiarizes students with the research publication process. This process not only stimulates faculty student collaboration and provides students with valuable feedback on their research, but also promotes collaboration between the College and Harvard's many graduate and professional schools. In addition to publishing original student research papers, THURJ keeps the Harvard community updated on and provides an important forum for discourse on the cutting-edge research that impacts our world today.

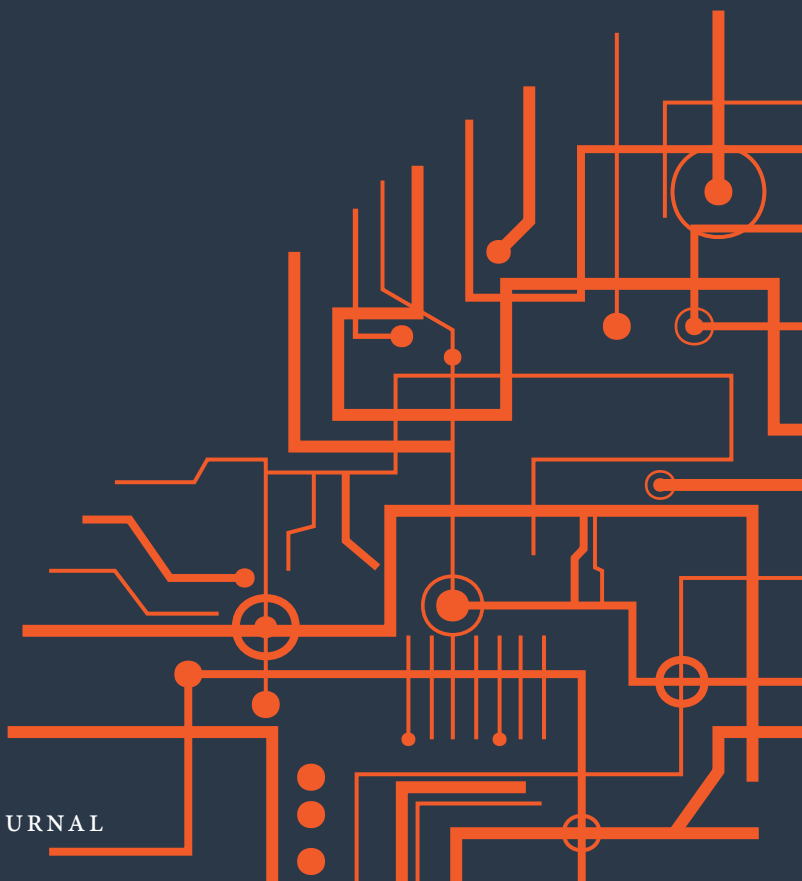
About the Cover

Iris Sung '27

The cover explores the ways in which the quest for knowledge evolves through time. Courses were once charted on physical maps, by compass and ship. Although we cannot abandon them completely, the ways in which we accumulate, and even understand, knowledge have become less tangible and concrete, and much more convoluted. The texture of the orange network ranges from the organic (dragonfly wings and river systems), to the artificial (city maps and wired networks) to show that modern knowledge cannot always be clearly bifurcated between digital and organic, or even familiar and new.



Research



Comparative Analysis of Segmentation Models on the Human Visual Diet Dataset

Francesco Plastina
Harvard College '28

Teaching an algorithm to identify and separate objects within an image (segmentation) is central to AI applications, from delineating tumors in medical scans to safe robotics, augmented reality, and assistive navigation. Yet most benchmarks use web-style photos, whereas people learn from everyday, human-centric views. In this study, I ask whether training modern segmentation models on human-like images (Human Visual Diet dataset), scenes that resemble what people routinely encounter, changes models' performance and what these models attend to. I compare two conventional convolutional networks (ResNet-18 and DenseNet-121 with custom upsampling decoders) and a transformer-based SegFormer (pre-trained ViT) on their segmentation performance. To interpret model focus, I compute Gradient-weighted Class Activation Mapping (Grad-CAM) for the convolutional neural networks (CNNs) and inspect self-attention maps for the transformer (ViT). My results show that the ViT achieved the highest evaluation accuracy of ~80%, outperforming DenseNet (~74% accuracy) and ResNet (~57% accuracy). All models struggled with fine object boundaries and smaller objects, indicating room for improvement in edge localization. Grad-CAM for ResNet and DenseNet often appeared in diffuse or highlighted background regions, while the ViT attention maps were sometimes inconsistent, suggesting that these models do not consistently provide precise object degrees. While prior studies (e.g., Madan et al., 2024) suggest superior generalization for transformers, my HVD results address segmentation performance only: the transformer leads in accuracy, but all models remain below human-level segmentation, especially at fine boundaries and small objects. These findings underscore the difficulty of achieving human-level generalization in segmentation, aligning with recent work emphasizing the importance of more human-like data 'diets' and architectures to close the gap between neural networks and human vision.

Introduction

Semantic segmentation, the process by which an algorithm labels every pixel in an image, supports applications such as tumor outlining in medical scans, robotics, and assistive navigation. A central challenge in segmentation is generalization: keeping high accuracy when the viewpoint, lighting, or context changes. Most benchmarks use web-style, object-centric photos that miss much of everyday visual variation. The Human Visual Diet (HVD) dataset addresses this by presenting objects in varied indoor scenes with controlled changes, more closely resembling what people routinely see (Madan et al., 2024).

This study tests how architecture affects segmentation on HVD and what image regions models rely on. I compare two convolutional neural networks (CNNs; ResNet-18 and DenseNet-121 with custom upsampling decoders) against a transformer-based SegFormer with a Vision Transformer (ViT) backbone. CNNs are a type of neural net that look at small patches of an image with the same tiny filter sliding around, so it can spot edges, textures, and shapes and build up to full objects (LeCun et al., 2015), while a transformer is a model that looks at all parts of the input image at once and decide which parts matter to each other using "attention" (Vaswani et al., 2017). To see where models "look," I use Grad-CAM, a method to explain a CNN's prediction by coloring the image to show which regions most influenced the decision, using gradients from the last convolution layer (Selvaraju et al., 2017), and attention maps, visualizations of the transformer's attention weights that indicate which patches it focused on, for the ViT. Results are reported as segmentation performance on a specific subset of HVD, without evaluating out-of-distribution generalization. In this framework, I compare the three architectures on HVD, examine training behavior and error patterns with an emphasis on small objects and fine boundaries, and relate accuracy differences to spatial focus using Grad-CAM and attention maps.

Prior work links computer-vision progress to findings in biological vision (Cox & Dean, 2014; Yamins et al., 2014), yet also shows important divergences: CNNs can rely on cues humans do not, and performance can drop under mild changes to images (Bowers et al., 2022; Linsley & Serre, 2023; Nagaraj et al., 2023; Serre, 2019). Vision transformers may capture object–context relations differently, but gains are not universal. Training on more realistic, human-like data has been proposed to improve alignment with human behavior (Nagaraj et al., 2023), and models trained on such "visual diets" can perform better under real-world transformations (Madan et al., 2024). Building on these observations, I use HVD to assess how architectural choices shape segmentation performance and whether higher accuracy coincides with more appropriate spatial focus.

Methods

Dataset and Preprocessing

I used the Human Visual Diet (HVD) dataset introduced by Madan et al. (2024), focusing on the photorealistic renderings of indoor scenes contained within the *main_xml* folder. Each image depicts a target object in a fully furnished room with varied layouts, materials, and lighting. Ground-truth segmentation labels were provided as .npy masks in *labels_main_xml*. Each pixel is labeled with an integer class ID corresponding to an object category (e.g., bed, sofa, table, wall, floor, etc.). There are 27 distinct class IDs present in my subset of HVD (1, 4, 5, 10, 11, 13, 21, 23, 31, 32, 37, 41, 42, 43, 44, 45). The object each class corresponds to can be found in the dataset folder. For training,

I applied minimal preprocessing: images were resized to 64 × 64 pixels (to reduce memory, as a compromise given computational constraints) and masks were likewise downsampled to 64 × 64 using nearest-neighbor interpolation. The pixel intensity values were normalized to [0, 1]. While this downsampling sacrificed detail (i.e., edges become less crisp), it allowed training of multiple models within a reasonable time with low computational power. I randomly split the dataset into 80% training and 20% validation images. No test set of completely novel scenes was used; the evaluation focused on the validation performance for generalization to unseen images of similar distribution. Because my evaluation was in-dataset (no novel-scene test), I kept the training setup the same for every model. This ensured a fair basis for comparison: any differences stemmed from the architectures, not from training choices.

Model Architectures

Three segmentation models were implemented:

1. **ResNet-18 + Upsampling Decoder:** Using a ResNet-18 backbone pretrained on ImageNet (up to the final convolutional block, removing the classification head). This provided a 512-channel feature map (of size 2 × 2 after my 64 × 64 input), which was fed into a custom decoder, a series of 2 × 2 bilinear upsamples with intervening 3 × 3 convolutional layers reducing channels. The final layer was a 1 × 1 convolutional layer producing class logits for each of the 27 classes, and I upsampled this to the original image size.
2. **DenseNet-121 + Upsampling Decoder:** Similarly, I took DenseNet-121 pretrained (on ImageNet) and truncated at its convolutional feature map. A decoder with three upsample steps (and convolutional layers halving channels each step) was used, yielding final class logits at 128 × 128, which were then interpolated to 128 × 128 (my chosen output size for DenseNet model). The ResNet and DenseNet decoders were implemented from scratch, but following a typical fully convolutional network design for segmentation.
3. **SegFormer (ViT):** A vision transformer-based segmentation model. I used the HuggingFace Transformers library's SegformerForSemanticSegmentation implementation, starting from the pre-trained weights on ADE-20k. This model has an efficient MiT-B0 backbone and a lightweight MLP decoder. I modified the final classification head to produce 27 classes instead of ADE-20k's 150, and allowed the head weights to initialize randomly.

For fair comparison, I fine-tuned all models on the HVD training set.

Training Procedure

I trained ResNet and DenseNet for up to 100 epochs, while SegFormer was fine-tuned for 50 epochs as it converged faster. I employed cross-entropy loss on the pixel labels, ignoring unlabeled pixels (if any, since most pixels were labeled). During training, I tracked the pixel-wise accuracy (percentage of pixels correctly classified) across classes on both training and validation sets. Pixel accuracy is a coarse measure that can be high if large areas are correctly labeled, even if smaller objects are not correctly classified. No heavy data augmentation was applied, given that the dataset already contained substantial variation.

Grad-CAM Implementation

To interpret the CNN models used in this project, specifically for ResNet and DenseNet, I implemented Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017). This method highlights the regions of an input image that most strongly influence the model's decision for a particular class. The process begins by passing an image through the network to obtain the prediction scores for each class. For a chosen class, I computed the gradient of its score with respect to the feature maps of the last convolutional layer in the encoder. These gradients indicate how sensitive the class score is to changes in different spatial regions of the feature maps. By averaging the gradients across the spatial dimensions, I obtained a single weight for each channel in the feature map, reflecting its importance to the class prediction. Each channel was then multiplied by its corresponding weight, and the results were summed to create a raw heatmap. Applying a ReLU function ensures that only positive contributions are kept, emphasizing the regions that positively influence the prediction. Finally, the heatmap was upsampled to the size of the input image and overlaid using a color map for visualization. This technique visualized where the model was focusing when identifying objects, and assessed whether it was attending to relevant parts of the image or being distracted by background features. This method was developed by Selvaraju et al. (2017).

Transformer Attention Maps

To interpret the SegFormer model's predictions, I extracted its internal self-attention matrices to visualize where the model focuses when making decisions for a given class. SegFormer, unlike some transformers, doesn't use a special classification token. Instead, it processes the input image as a grid of patches. For the B0 variant I used, the image was downsampled by a factor of 32 in height and width so that the final attention maps operate on a much smaller patch grid. During evaluation, I enabled the model to return its attention weights, which are square matrices showing how much each patch in the image attends to all other patches.

To generate attention maps for a specific class, I first ran a forward pass to obtain the predicted segmentation mask and the attention weights from the final transformer layer. I then identified the set of patch positions that were labeled as the class of interest, such as "lamp" or "bed." For each of these patches, I extracted its corresponding row in the attention matrix, which represents how strongly that patch attends to every other patch. I averaged these attention vectors across all selected patches and across all attention heads to get a single, representative attention map for that class. Then, I reshaped this one-dimensional attention vector into a 2D grid matching the layout of the patch map, and upsampled it to the size of the original input image. Finally, I overlaid the resulting heatmap on the image to visualize where the model was attending when making its class prediction. This attention visualization doesn't directly explain what caused the prediction, like Grad-CAM, but it shows which regions the model relied on when assigning a class to specific parts of the image. I applied this technique to two different class predictions to study how the transformer uses spatial context. My implementation was based on the HuggingFace SegFormer model's output, and while it was inspired by more advanced attention rollout

methods, I kept the approach simple and focused on per-class attention accumulation.

Comparison

I directly compared the models on their pixel accuracy performance to evaluate their segmentation capabilities, estimating their IoUs (Intersection over Union) that describe how much two regions (the ground truth and the prediction of the model) overlap. Moreover, I compared how the different models decide on assigning a specific pixel class by directly showing their Grad-CAM (for ResNet and DenseNet) and attention (for ViT) maps.

Results

Segmentation Performance

All three models successfully learned to segment the HVD scenes to a reasonable degree, but with notable differences in accuracy. The Vision Transformer (SegFormer-B0) achieved the best overall performance (Fig. 1), followed by DenseNet (Fig. 2), then ResNet (Fig. 3). On the validation set, the SegFormer attained pixel accuracy of approximately 81% and an estimated mean IoU of around 0.60. DenseNet's accuracy was about 72% at convergence (mean IoU ~ 0.50), and ResNet's was around 57% (mean IoU ~ 0.35–0.40). These quantitative gaps reflect the models' capacity differences and possibly the benefit of SegFormer's pretrained features. Training and validation loss curves showed that ResNet-18 underfitted the data: its validation accuracy peaked early and then stagnated, indicating it struggled with the complexity of the task (many classes, high variability), given its smaller capacity. DenseNet-121, with its deeper architecture, learned faster and attained higher accuracy, but eventually showed signs of overfitting, as validation loss began rising after about 90 epochs while training loss kept decreasing. The SegFormer converged rapidly, within about 10–15 epochs, to a low validation loss, suggesting that the pretrained transformer features were already well-aligned to the segmentation task, especially since ADE-20k pre-training provides a strong prior on indoor scenes.

Training Dynamics

Figures 1–3 illustrate the validation pixel accuracy curves for the models. ResNet's curve starts lowest and improves slowly, leveling off around 55–58%. DenseNet starts higher and climbs to approximately 74% before slightly dipping (because of overfitting). SegFormer starts around 62% (already high due to pretraining) and improves to about 81%. The relative ordering of the curves stayed consistent across training. Validation accuracy and training accuracy followed similar trends, though training accuracy could reach above 90% for DenseNet and SegFormer, indicating some overfitting (the gap between training and validation accuracy for DenseNet was about 20% by epoch 100). I attempted mild regularization (dropout in the decoder, weight decay) for DenseNet, which narrowed this gap but did not fundamentally change the final results. The transformer, despite its capacity, did not severely overfit, likely because the pretrained weights acted as a strong regularizer.

Qualitative Segmentation Examples

I inspected segmentation outputs on several validation images to understand the models' successes and failures. In general, SegFormer produced the most coherent and correct segmentations, although it showed some distorted overlays with the input images. DenseNet also performed surprisingly well, as its prediction mask did not differ greatly from the ground truth. ResNet performed poorly, though it was still able to segment between walls, floor, and other objects. Each model had significant difficulty in recognizing edges and boundaries, and in general, the 3D structure of the images. Figures 4-6 show a sample of the same input image and the outputs of each model.

Grad-CAM Visualization (CNNs)

I applied Grad-CAM to interpret where ResNet-18 and DenseNet-121 focus their attention for predicting certain classes. For the target object class in each scene—the object the model was asked to identify in context—Grad-CAM maps tended to highlight broad regions that only partially overlapped the object. For example, in a living room scene where the target is "floor," the Grad-CAM for class "floor" on DenseNet highlighted a specific side of the floor (Fig. 7), and highlighted part of the wall and of an object nearby. This suggests the model's prediction of "wall" was influenced by features in adjacent regions or simply due to imprecise localization in the feature maps. ResNet's Grad-CAM for the same class in other images was even more spread out, essentially covering the entire image (Fig. 8). Overall, Grad-CAMs for DenseNet were slightly more localized than ResNet's, reflecting DenseNet's superior segmentation performance. Yet in comparison to humans, who would squarely focus on the object of interest, the Grad-CAMs appeared noisy and overinclusive. This aligns with known issues that CNNs don't always have interpretable internal representations, but they may rely on context textures.

Transformer Attention Visualization

For the SegFormer model, I examined its last-layer attention maps for selected classes. While I expected the transformer's attention mechanism to show a more structured pattern, I instead found largely unstructured attention maps. For example, two very different classes, "cabinet" (42) and "floor" (44) shared similar attention maps and did not consider the "floor" at all (Fig. 9). In summary, the ViT's attention maps clearly did not align with human-like rationale, which tends to focus the attention on the object for recognition. This mirrors the Grad-CAM issue: higher accuracy didn't necessarily translate into more pinpointed attention. It seems the SegFormer achieves higher accuracy through its ability to integrate features over a larger receptive field, but it still spreads its attention across the scene in ways that are hard to interpret.

Discussion

Practical Significance

On the HVD indoor scenes, a compact transformer (SegFormer) achieved the highest in-dataset pixel accuracy (approximately 80%) relative to two CNN baselines (DenseNet: approximately 74%, ResNet: approximately 57%). For professionals working with HVD-like data and limited compute, this identifies a strong default baseline. At the same time, all models showed consistent errors on small objects and along fine boundaries. Therefore, in applications where edge precision matters (e.g., medical delineation, manipulation in robotics), these failure modes will directly limit utility unless addressed.

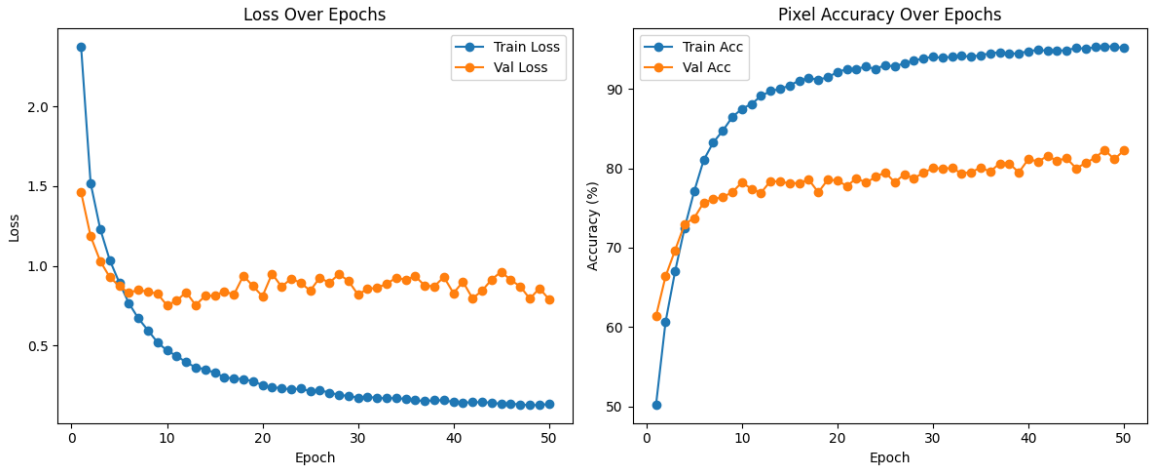


Figure 1. *ViT Segmentation Performance:* The graph on the left shows ViT's training and the validation loss. Since the training loss (blue) is going down while the validation loss (orange) is almost constant, also ViT is overfitting, but less than the previous two models. Indeed, despite the pixel accuracy of the model on the training data (blue) is higher than the accuracy of the model on the validation data (orange), the two graphs are closer than before.

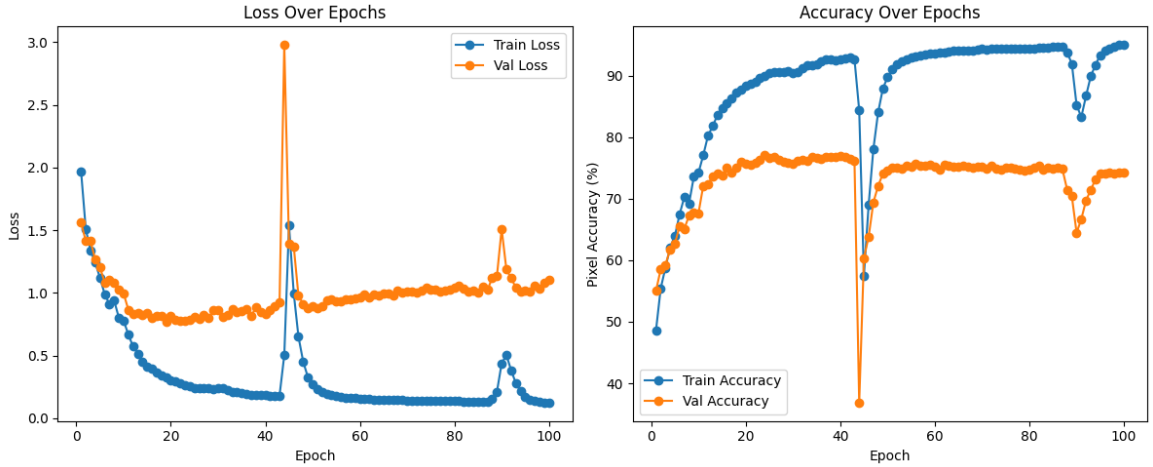


Figure 2. *DenseNet Segmentation Performance:* The graph on the left shows DenseNet's training and the validation loss as explained in Figure 1. Since the training loss is going down and the validation loss is going slightly up, DenseNet is aslo overfitting, but less than ResNet. The overfitting is more clear in the figure on the left where the pixel accuracy of the model on the training data (blue) is higher than the accuracy of the model on the validation data (orange), while still better than ResNet performance.

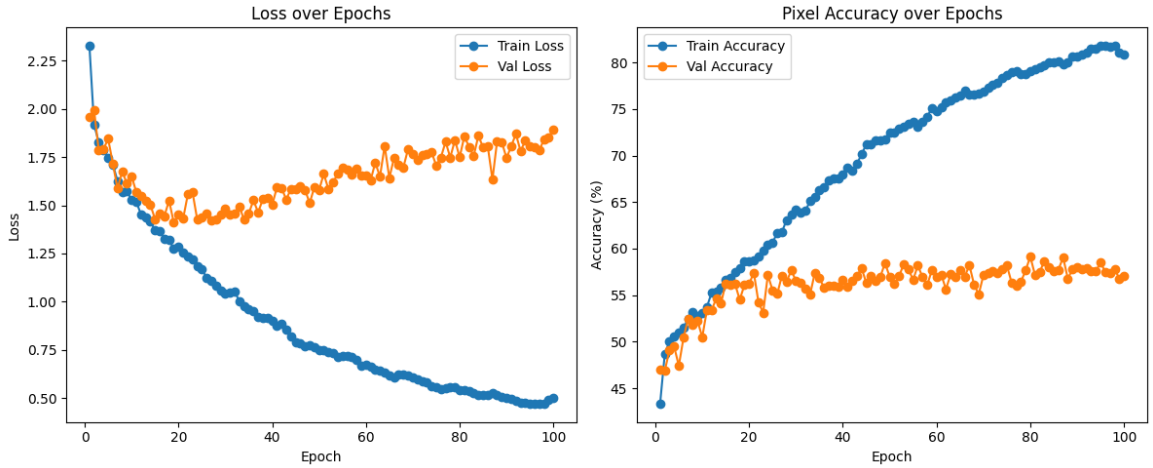


Figure 3. *ResNet Segmentation Performance:* The graph on the left shows ResNet's training and the validation loss. The training loss (blue) is measured on the training data while the model is learning, while the validation loss (orange) is measured on a separate hold-out set that the model never learns on and therefore estimates how well the model will do on new data. Since the training loss is going down while the validation loss is going up, ResNet is overfitting, which means that the model learns the training data so well that it doesn't generalize to new data. The overfitting is more clear in the figure on the left where the pixel accuracy of the model on the training data (blue) is way higher than the accuracy of the model on the validation data (orange).

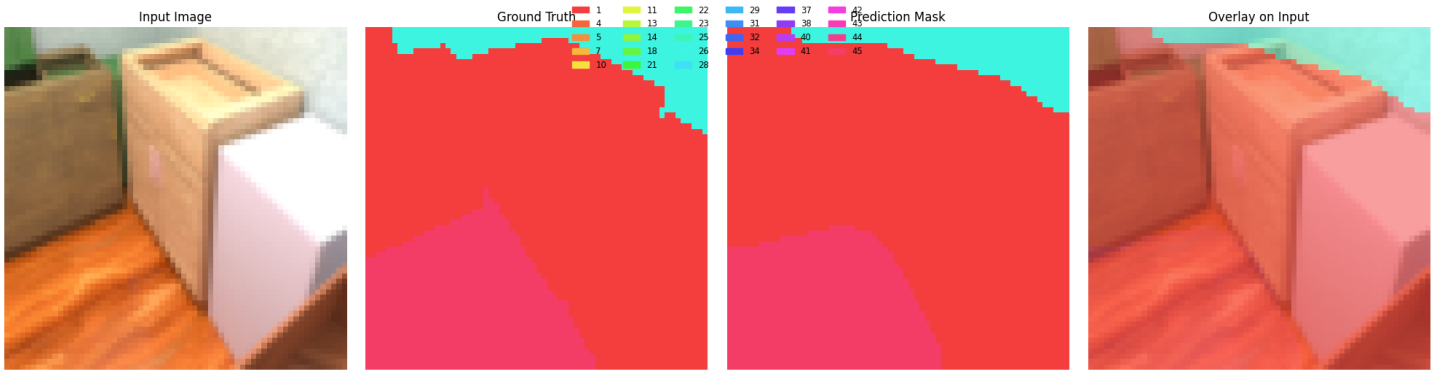


Figure 4. *ResNet output sample:* These images show a sample of ResNet's output prediction and compare it with the input image. Starting from the left, in order it is possible to see the input image, the ground truth, the prediction mask of the model (its output), and the overlay of the prediction on the input image. For each color corresponds one of the 45 different classes of the dataset.

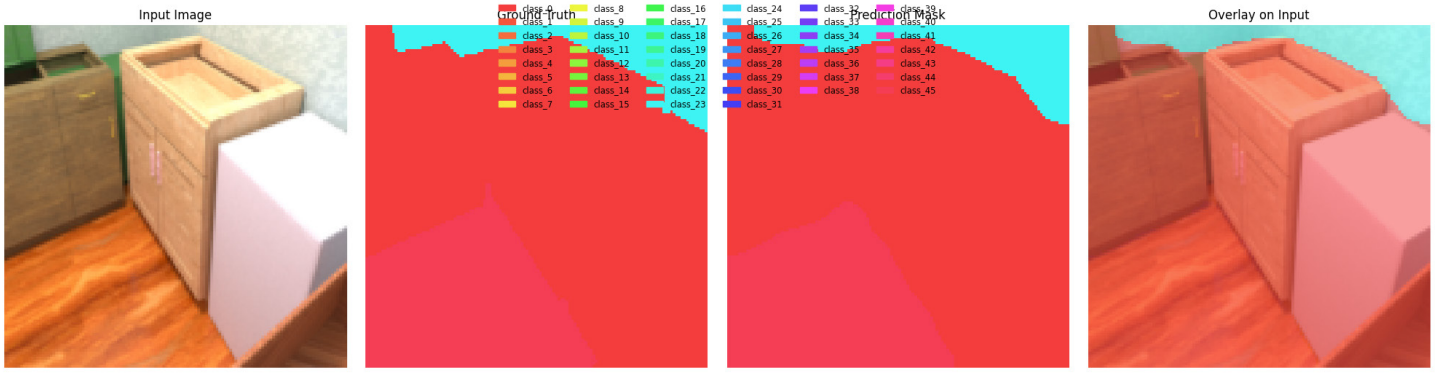


Figure 5. *DenseNet output sample:* These images show a sample of DenseNet's output prediction and compare it with the input image. Starting from the left, in order it is possible to see the input image, the ground truth, the prediction mask of the model (its output), and the overlay of the prediction on the input image. For each color corresponds one of the 45 different classes of the dataset.

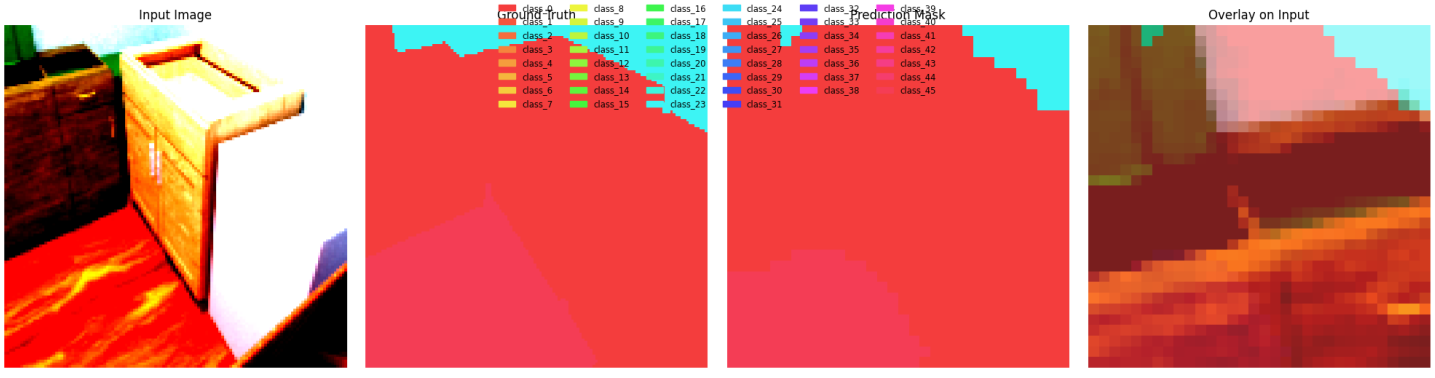


Figure 6. *ViT output sample:* These images show a sample of ViT's output prediction and compare it with the input image. Starting from the left, in order it is possible to see the input image, the ground truth, the prediction mask of the model (its output), and the overlay of the prediction on the input image. For each color corresponds one of the 45 different classes of the dataset.

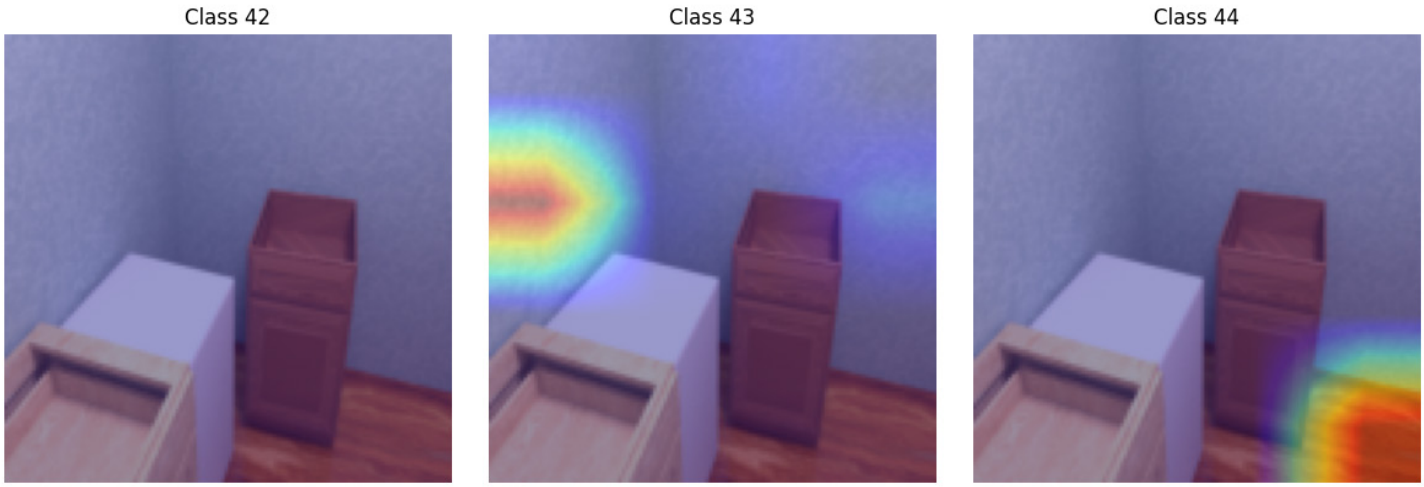


Figure 7. *DenseNet Grad-Cam visualization sample:* These images show which elements of the image DenseNet focused on to make the segmentation decision (assigning the pixels to their specific class) for three random classes (42, 43, 44). Warm colors correspond to higher "focus," while cool colors to less "focus."

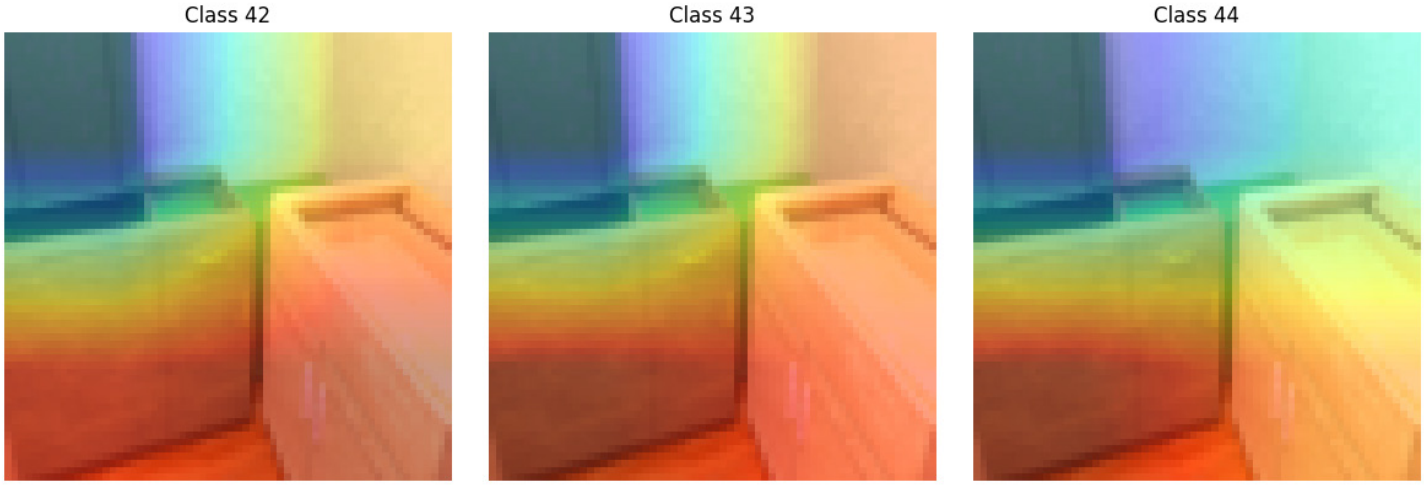


Figure 8. *ResNet Grad-Cam visualization sample:* These images show which elements of the image ResNet focused on to make the segmentation decision (assigning the pixels to their specific class) for three random classes (42, 43, 44). Warm colors correspond to higher "focus," while cool colors to less "focus."

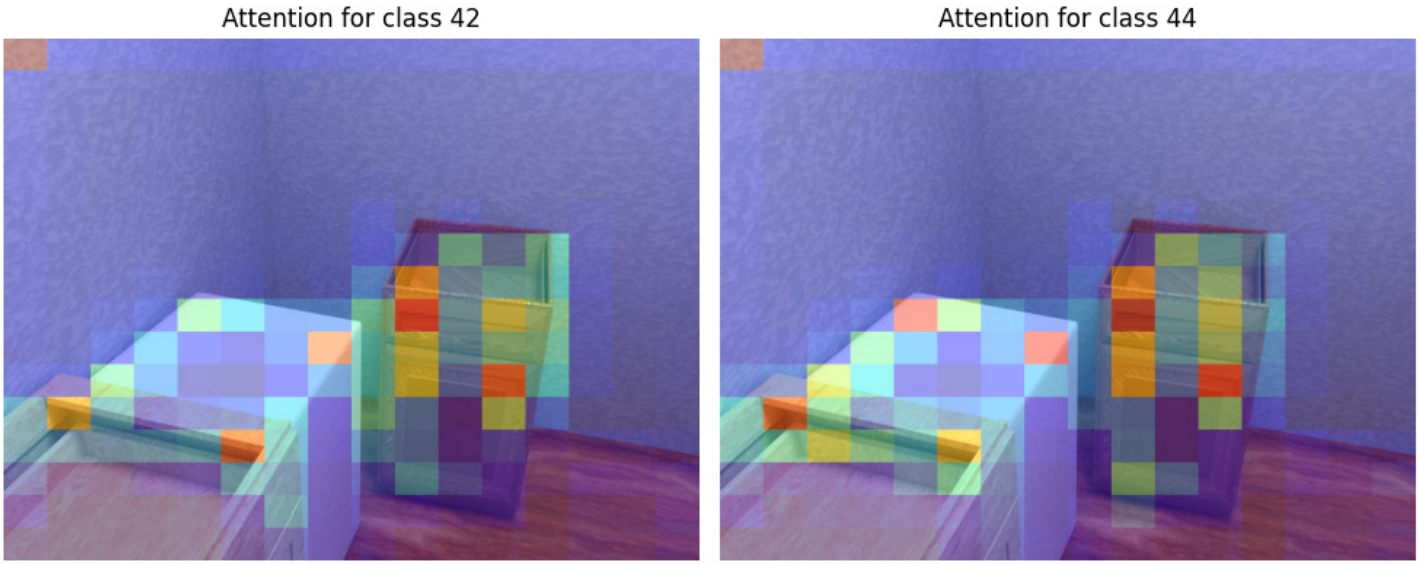


Figure 9. *ViT attention maps:* These images show which elements of the image ViT focused its "attention" on to make the segmentation decision (assigning the pixels to their specific class) for two random classes (42, 44). Warm colors correspond to higher "focus," while cool colors correspond to less "focus."

Accuracy vs Human-like Reasoning

Qualitative analyses (Grad-CAM for CNNs, attention maps for the transformer) did not reveal reliable, object-tight focus. CNN heatmaps often spread into the background; transformer attention was instead broad or inconsistent. These analyses caution against interpreting higher accuracy as evidence of human-like spatial reasoning and suggest that current explanation maps are, at best, sanity checks rather than proof of mechanism.

Context: Necessary but Insufficient

HVD contains natural variation in viewpoint, lighting, and materials, and the models maintained in-dataset performance across these variations in the validation split. The results, therefore, support using human-like data for evaluation, but also indicate that data realism alone does not fix boundary or occlusion errors.

Takeaway

Within HVD, transformers offer a practical accuracy advantage under matched (or near-matched) conditions, but all tested models remain limited by boundary and small-object errors and do not show reliably human-like spatial focus. Closing this gap likely requires pairing human-like data with inductive biases and objectives that explicitly encode edges, scale, and occlusion, not just scaling existing architectures or relying on generic pretraining.

Limitations

Pre-training and Resolution Mismatches

To keep the study reproducible and focused, I used standard pre-trained features for each network (ImageNet for CNNs; ADE20K for SegFormer with a reinitialized 27-class head) and matched evaluation on a common 128×128 grid despite differing training resolutions between ResNet and the other networks due to GPU limits. These choices favor practical, commonly used configurations over custom, model-specific tuning. As a result, I limit claims to in-dataset performance; I do not claim out-of-distribution generalization, and I note that resolution and pretraining differences may contribute to the observed gaps. Qualitative analyses (Grad-CAM and attention maps) and a discussion of small-object/boundary errors are included to make these trade-offs explicit.

Downsampling

The study was made with a limited computational power and memory using the free time-limited GPUs offered on Google Colab. The large size of the dataset and of the images was impossible to handle with such limited computational power. To tackle this challenge, I decided to downsample. While I recognize that downsampling may create a loss of some information, since these important results were shown on a smaller scale, it is reasonable to assume that they will be shown on the larger data.

Grad-CAM and Attention Maps

The use of Grad-CAM and attention maps in the study aims to be only a qualitative analysis of how these models are making their decisions while assigning each class to each pixel. Further studies may focus on follow-up analysis to quantify the attention/feature maps.

Conclusions

On the Human Visual Diet dataset, a compact transformer (SegFormer-B0) outperformed two CNN baselines, yet none approached human-level segmentation, with persistent errors on fine boundaries and small objects. Interpretability analyses (using Grad-CAM for CNNs and attention maps for the transformer) revealed diffuse or inconsistent focus, cautioning against equating higher accuracy with human-like reasoning. These results suggest that human-like data are helpful but insufficient. Closing the gap will require objectives and architectures that explicitly encode edges, scale, and occlusion, alongside higher-resolution training and evaluation beyond in-dataset splits. Future work should probe out-of-distribution generalization to novel scenes and explore edge-aware losses or multi-scale decoders.

Code Availability

I built the training pipeline and custom decoders from scratch in Google Colab for ResNet and Densenet models, while the SegFormer model utilization relied on the HuggingFace Transformers library implementation. The code for this project is available at https://github.com/fran0324/Neuro-240-Project.

References

Bowers, J. S., Malhotra, G., Dujmovic, M., Montero, M. L., Tsvetkov, C., Biscione, V., Puebla, G., Adolfi, F., Hummel, J. E., Heaton, R. F., Evans, B. D., Mitchell, J., & Blything, R. (2022). *Deep Problems with Neural Network Models of Human Vision. The Behavioral and Brain Sciences*, *46*, e385.

Cox, D. D., & Dean, T. (2014). *Neural Networks and Neuroscience-Inspired Computer Vision. Current Biology*, *24*(18), 2073-2216.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep Learning. Nature*, *521*(7553), 436–444.

Linsley, D., & Serre, T. (2023). *Fixing the problems of deep neural networks will require better training data and learning algorithms. The Behavioral and Brain Sciences*, *46*, e400.

Madan, S., Li, Y., Zhang, M., Pfister, H., & Kreiman, G. (2024). *Improving out-of-distribution generalization by mimicking the human visual diet. "NeuroAI: Fusing Neuroscience and AI for Intelligent Solutions" NeurIPS workshop.*

Nagaraj, A., Ashok, A. K., Linsley, D., Lewis, F. E., Zhou, P., & Serre, T. (2023). *Ecological data and objectives align deep neural network representations with humans. "UniReps: Unifying Representations in Neural Models" NeurIPS workshop.*

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). *Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. IEEE International Conference on Computer Vision (ICCV)*, 618-626.

Serre, T. (2019). *Deep Learning: The Good, the Bad, and the Ugly. Annual Review of Vision Science*, *5*, 399–426.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need. Advances in Neural Information Processing Systems 30* (NIPS 2017).

Yamins, D. L., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). *Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8619–8624.

Cryptological Applications of Probability Theory

Mohamed Humaid Saleem
Harvard College '26

There exists a deeply intertwined relation between cryptography and probability theory. Most fundamental concepts in cryptography, like cryptographic security, are based on probabilistic reasoning. This paper will explore concepts related to Chernoff bounds in cryptanalysis, pseudo-random number generation, entropy and information-theoretic security, and statistical randomness tests, through the use of rigorous definitions, theorems, and proofs. We conclude by exploring a historical case study of Andrew Gleason's tail-bound attack on the Japanese "Coral" cipher during World War II, as well as some modern-day applications of probability theory in cryptography.

Introduction

The most concise description of cryptology is that it is the science of secure communication. Communication is secure only if an adversary has a negligible probability of deciphering messages; thus, cryptology is fundamentally probabilistic. Cryptanalysis utilizes statistical tests to decipher non-random patterns in ciphertexts. Concepts in probability theory such as entropy and randomness are used to quantify uncertainty in keys, messages, and adversarial behavior. This paper will analyze the use of such probabilistic techniques. Section 2 introduces entropy and Shannon's perfect secrecy. Section 3 explores pseudorandom number generators (PRNGs). Section 4 discusses Chernoff bounds, with a WWII case study on breaking the Japanese Coral cipher. Section 5 covers probabilistic encryption. Section 6 examines statistical tests for randomness.

We will first define cryptographic jargon that will be used frequently throughout the paper:

Definition 1.1 (Cryptanalysis). Cryptanalysis is the science of breaking codes and deciphering encrypted messages without knowing the secret key.

Definition 1.2 (Encryption and Decryption). Encryption is the process by which a readable message is converted to an unreadable form to prevent unauthorized parties from reading it. Decryption is the process of converting an encrypted message back to its original (readable) format.

Definition 1.3 (Keys). Cryptographic keys are classified into two main types: symmetric and asymmetric. Symmetric keys use a single key for both encryption and decryption, while asymmetric keys use two mathematically related keys: a public key for encryption and a private key for decryption.

Definition 1.4 (Ciphertexts). Ciphertext is the result of applying a cipher (an encryption algorithm) to plain text to make it unreadable without the proper decryption method.

Entropy and Information-Theoretic Security

Entropy measures uncertainty in cryptographic keys or messages, underpinning information-theoretic security. We will explore in depth what it means for a message or key to be uncertain, in relation to entropy.

Definition 2.1 (Shannon Entropy). Let X be a discrete random variable with values $x_1,\ldots, x_n$ and probabilities $p_i = P[X = x_i]$. The *entropy* is

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i.$$

Remark 2.1. Within the realm of information theory, entropy is defined using the base-2 logarithm as we measure uncertainty in bits. However, we shall use the natural logarithm to achieve generality. To do this, we need to adjust the unit of measurement to nats (bits for base-2, nats for base-e).

For a key K uniformly chosen from N possibilities, $H(K) = \log N$. High entropy strengthens keys by increasing guessing difficulty. For example, consider a 4-bit key, K with probabilities $p(0000) = 1/2$, $p(0001) = 1/4$, $p(0010) = 1/8$, $p(0011) = 1/8$, and $p(\text{others}) = 0$. The entropy is:

$$H(K) = -\left(\tfrac{1}{2}\log\tfrac{1}{2} + \tfrac{1}{4}\log\tfrac{1}{4} + \tfrac{1}{8}\log\tfrac{1}{8} + \tfrac{1}{8}\log\tfrac{1}{8}\right) = \tfrac{1}{2}\cdot\log 2 + \tfrac{1}{4}\cdot\log 4 + \tfrac{1}{8}\cdot\log 8 + \tfrac{1}{8}\cdot\log 8.$$

Using $\log 2 \approx 0.693$, $\log 4 \approx 1.386$, $\log 8 \approx 2.079$, we get:

$$H(K) \approx \tfrac{1}{2}\cdot 0.693 + \tfrac{1}{4}\cdot 1.386 + \tfrac{1}{8}\cdot 2.079 + \tfrac{1}{8}\cdot 2.079 \approx 0.347 + 0.347 + 0.269 + 0.260$$
$$= 1.214 \text{ nats.}$$

Observe that for a uniform 4-bit key:

$$H(K) = \log 16 \approx 2.773 \text{ nats.}$$

Then our previously calculated entropy is low in comparison, and thus the key is vulnerable to being deciphered by an adversary, as the key is biased toward certain values. Therefore, cryptographic keys aim for maximum entropy to resist attacks (Menezes et al., 1997). Essentially, we are trying to ensure that the ciphertext does not reveal the original, unencrypted message (plain text). One such way of doing this is Shannon's perfect secrecy.

Definition 2.2 (Perfect Secrecy). Let

$$M : \Omega \to \mathscr{M}, \quad K : \Omega \to \mathscr{K}, \quad C : \Omega \to \mathscr{C}$$

be the plaintext, key, and ciphertext random variables on a probability

space (Ω, P). Assume K is chosen uniformly over $\mathscr{K}$ and $C = E(M, K)$ for some encryption function/scheme $E : \mathscr{M} \times \mathscr{K} \to \mathscr{C}$. The scheme has *perfect secrecy* if for every $m \in \mathscr{M}$ and every $c \in \mathscr{C}$ with $P[C = c] > 0$,

$$P[M = m \mid C = c] = P[M = m].$$

Equivalently, M and C are statistically independent:

$$P[M = m \wedge C = c] = P[M = m]\ P[C = c].$$

Remark 2.2. $M = m$ denotes the event

$$\{\omega \in \Omega \mid M(\omega) = m\}.$$

We write $\mathscr{M}$ for the message space (the set of all possible plaintexts m), $\mathscr{K}$ for the key space (the set of all possible secret keys k, and $\mathscr{C}$ for the ciphertext space (the set of all possible ciphertexts c).

Shannon's theorem proved that in order to achieve perfect secrecy, the entropy of the key must be equal to or exceed that of the message.

Theorem 2.1 (Shannon Theorem). For perfect secrecy within a finite message space, the size of the keyspace must be at least as large as that of the message space, mathematically speaking, $H(K) \geq H(M)$.

The proof of this theorem in Shannon is quite long (1949); for the sake of brevity, we only provide an overview below:

Proof. Perfect secrecy requires $P[C = c \mid P = m_1] = P[C = c \mid P = m_2]$ for all plaintexts m_1, m_2 and ciphertext c. Each plaintext must have the same number of keys mapping to each ciphertext, implying $|\mathscr{K}| \geq |\mathscr{M}|$. Thus, $H(K) \geq H(M)$ (Shannon, 1949). □

Remark 2.3. Perfect secrecy is not pragmatic as Shannon's theorem dictates that very large messages require the use of large, one-time use keys, which are hard and cost-ineffective to create. However the concept of entropy is still highly valuable as it informs computational security. For example, a 128-bit key with $H(K) = \log(2^{128}) \approx 88.7$ nats requires 2^{128} guesses, infeasible for adversaries. Entropy also allows us to quantify breaches such as protocol leakage, thus guiding secure design (Menezes et al., 1997).

Pseudorandom Number Generation

Almost all cryptographic systems require random bits in order to generate keys; oftentimes, this is done through pseudorandom number generators (PRNGs).

Definition 3.1 (PNRG). A PRNG $G : \{0,1\}^s \to \{0,1\}^n$ ($n >> s$) expands a random seed (an n-bit string chosen at random) into a string that cannot be differentiated from random bits. Cryptographically secure PRNGs make it such that no efficient algorithm can distinguish $G(U)$ from a random n-bit string (Blum et al., 1986), where U is some random variable distributed uniformly over $\{0,1\}^s$.

The Blum–Blum–Shub (BBS) generator is a great example of a PRNG in cryptography.

Definition 3.2 (Blum-Blum-Shub PRNG). Choose 2 primes p, q such that $p, q \equiv 3 \pmod 4$. Let $N = pq$. Let x_0 be a seed that is coprime to N, and iterate such that $x_{i+1} \equiv x_i^2 \pmod N$. Finally each output is the least significant bit of each x_i, and the subsequent string is cryptographically secure as desired.

Before ending this section, we shall make a quick note on why BBS is secure, despite being somewhat outdated. Its security stems from the fact that deciphering every output bit is dependent on how difficult it is to factor $N = pq$. In practice, BBS sets p and q such that they are large primes, making such a factorisation quite hard. Furthermore, each iteration squares some number modulo N, which produces a quadratic residue (a number z for which there exists an x such that $x^2 \equiv z \pmod N$) and outputs the least significant bit. Predicting each of these bits requires an adversary to differentiate quadratic residues, a problem that is regarded as hard as factoring N, which is widely considered computationally infeasible. Thus, BBS ensures pseudorandomness (Blum et al., 1986; Menezes et al., 1997).

Chernoff Bounds and Cryptographic Analysis

Chernoff bounds are useful is cryptology as they give us exponentially decreasing tail probabilities, which are ideal for identifying deterministic cipher behavior in cryptoanalysis. There are many types of Chernoff bounds, each of which has its own set of assumptions. We will discuss the Chernoff bound for the simple case of a sum of independent Bernoulli trials, where each random variableakes values 0 or 1. A great example is tossing unfair coins, each with its own probability of heads, and counting the total number of heads.

Theorem 4.1 (Chernoff Bound). Let $X = \sum_{i=1}^{n} X_i$, where $X_i = 1$ with probability p_i and $X_i = 0$ with probability $1 - p_i$, and all X_i are independent. Let $\mu = \mathbb{E}[X] = \sum_{i=1}^{n} p_i$. Then:

(i) Upper Tail: $P[X \geq (1 + \delta)\mu] \leq e^{(-\frac{\delta^2\mu}{2+\delta})}$ for all $\delta > 0$,
(ii) Lower Tail: $P[X \leq (1 - \delta)\mu] \leq e^{(-\frac{\mu\delta^2}{2})}$ for all $0 < \delta < 1$.

Proof. We begin by recalling Markov's Inequality:

Lemma 4.1 (Markov's Inequality). Let $X : S \to \mathbb{R}$ be a non-negative random variable. Then, for any $a > 0$,

$$P[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$

We can now proceed with proving Theorem 4.1, utilizing Lemma 4.1. For the upper tail, we manipulate Markov's inequality by using exponentiation. Observe that for $s > 0$, we have that

$$X \geq (1 + \delta)\mu \implies e^{sX} \geq e^{s(1+\delta)\mu}.$$

Using the above and applying lemma 4.1, we get:

$$P[X \geq (1+\delta)\mu] \leq \frac{\mathbb{E}[e^{sX}]}{e^{s(1+\delta)\mu}}.$$

Independence of X_i gives

$$\mathbb{E}[e^{sX}] = \prod_{i=1}^{n}(1 - p_i + p_i e^s) \leq e^{((e^s-1)\mu)}$$

since $1 - p_i + p_i e^s \leq e^{(p_i(e^s-1))}$. Substituting $s = \log(1 + \delta)$ yields

$$\mathbb{E}[e^{sX}] \leq e^{(\delta\mu)},\ e^{(1+\delta)\mu\log(1+\delta)} = (1+\delta)^{(1+\delta)\mu}.$$

Therefore,

$$P\left[X \geq (1+\delta)\mu\right] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{\mu}.$$

Now observe that taking the logarithm of the right hand side, in the above equation, still preserves our inequality. Taking the logarithm of the RHS yields $\mu(\delta - (1 + \delta)\log(1 + \delta))$.Then using $\log(1 + \delta) \geq \delta/(1 + \delta/2)$ we get $(1 + \delta)\log(1 + \delta) \geq \frac{\delta(1+\delta)}{1+\delta/2}$, so $\delta - (1 + \delta)\log(1 + \delta) \leq \delta - \frac{\delta(1+\delta)}{1+\delta/2} = \frac{\delta(1+\delta/2)-\delta(1+\delta)}{1+\delta/2} = -\frac{\delta^2}{2+\delta}$. Finally recall that $\forall x \in \mathbb{R}$ *s.t.* $x < 0$, we have that $e^x > x$, applying this to our previous result yields the desired equation,

$$P[X \geq (1 + \delta)\mu] \leq e^{\left(-\frac{\delta^2\mu}{2+\delta}\right)}.$$

The proof of the lower tail is analagous, with the only exceptions being the use of the substitution $s = \log(1 - \delta)$ and the result, $\log(1 - \delta) \geq -\delta + \delta^2/2$. □

Remark 4.1. We have attempted an efficient version of the proof using substitutions. A detailed sketch of this proof from first principles can be found in the lecture notes made by MIT lecturer Michel Goemans (2015).

Given we have proven the Chernoff bound, we can proceed to work out an example to illustrate how Chernoff bounds detect deviations from expected random behavior. Suppose we have a cipher with the standard 26-letter alphabet and we want to send a random ciphertext of length $n = 1000$. Then each letter should probabilistically have $1000/26 \approx 38.46$ occurrences due to randomness. Let $X_i = 1$ if the i-th letter is 'A', else $X_i = 0$, with $p_i = 1/26$ for all i. Then $X = \Sigma X_i$ counts 'A's, with $\mu = \Sigma p_i = 1000/26 \approx 38.46$. Suppose we want to know whether the occurrence of 60 or more 'A's is a significant deviation from the expected randomness of the ciphertext. We can test this using the upper tail bound. $t(1 + \delta)\mu = 60$, so $\delta = (60/38.46) - 1 \approx 0.56$. Then the probability of at least 60 'A's is,

$$P[X \geq 60] \leq e^{\left(-\frac{0.56^2 \cdot 38.46}{2+0.56}\right)} \approx e^{\left(-\frac{0.3136 \cdot 38.46}{2.56}\right)} \approx e^{(-4.71)} \approx 0.009.$$

This low probability indicates that observing 60 'A's is highly unlikely under a random ciphertext distribution, suggesting the cipher produces biased outputs. Such non-randomness reveals exploitable patterns in the cipher's structure, enabling an adversary to use cryptanalysis via frequency analysis (Section 6) or other statistical methods to decipher our message.

During WWII, Andrew Gleason utilized a variant of the Chernoff bound (a similar tail-bound, albeit pre-Chernoff), in order to break a highly sophisticated cipher known as the Japanese Coral cipher. He famously dubbed it the "Gleason Crutch". The Gleason Crutch analyzed ciphertext patterns to show that observed repetition within the Coral cipher was unlikely under randomness. This led to the revelation of the specific mechanics of the Coral cipher and thus led to its decryption in 1944 (Burroughs et al., 2009). In the modern age, Chernoff bounds are applied in side-channel attacks by analyzing deviations in power consumption or timing to infer key bits, ensuring negligible failure probabilities in cryptographic protocols.

Probabilistic Encryption Schemes

The primary motivation behind the discarding of deterministic encryption as a technique in cryptography is due to its tendency to leak plaintext patterns. We now prefer the use of probabilistic encryption as it allows for randomness, thus ensuring that identical plaintexts produce different ciphertexts, achieving semantic security. The Goldwasser–Micali scheme, a pioneering probabilistic encryption method, exemplifies this approach by utilizing the computational hardness of the quadratic residuosity problem.

Definition 5.1 (Key Generation for Goldwasser–Micali). To generate keys for the Goldwasser–Micali scheme, select two large prime numbers p and q, and compute $N = pq$. Define the multiplicative group $\mathbb{Z}^*_N = \{a \in \{1,..., N - 1\} \mid \gcd(a, N) = 1\}$, which is the set of integers under modulo N that are coprime to N. Choose $y \in \mathbb{Z}^*_N$ such that y is a *quadratic non-residue* under modulo N, meaning there exists no $x \in \mathbb{Z}^*_N$ satisfying $x^2 \equiv y \pmod N$, and the *Jacobi symbol* $(y/N) = +1$. The Jacobi symbol $(y/N) = \left(\frac{y}{p}\right)\left(\frac{y}{q}\right)$, where $\left(\frac{y}{p}\right)$ is the Legendre symbol (+1 if y is a quadratic residue modulo p, -1 if not, 0 if p divides y), ensures y is not trivially distinguishable from quadratic residues without knowledge of p and q.

Definition 5.2 (Encryption for Goldwasser–Micali). To encrypt a plaintext bit $b \in \{0, 1\}$ using the public key (N, y), select a random integer $r \in \mathbb{Z}^*_N = \{1,..., N\text{-}1 \mid \gcd(r, N) = 1\}$, chosen uniformly to ensure ciphertext unpredictability. Compute the ciphertext c as follows:

$$c = \begin{cases} r^2 \bmod (N) & \text{if } b = 0 \\ y \cdot r^2 \bmod (N) & \text{if } b = 1 \cdot \end{cases}$$

For $b = 0$, $c = r^2 \bmod (N)$ is a quadratic residue under modulo (N), since it is the square of r. For $b = 1$, $c = y \cdot r^2 \bmod (N)$ incorporates the quadratic non-residue y, making c a quadratic non-residue under modulo N, as y has no square root with respect to modulo (N). Therefore, the random choice of r ensures that every encryption of a given bit b creates a ciphertext that is unique and thus prevents the pattern leakage of the plaintext.

Statistical Tests for Randomness

Statistical tests validate randomness in PRNGs or ciphers, detecting biases that could enable attacks. We shall define 2 such statistical tests. Note that within this section, we shall not discuss in detail the specific applications of such tests. This is because their primary function is to test for randomness; thus, it suffices to prove that these tests provide a sufficient metric to evaluate randomness. We begin by defining the XOR operation in computer science, which will be relevant later in this section.

Definition 6.1 (XOR operation). The XOR operation, denoted $A \oplus B$, is a bit-wise binary function defined by:

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

In other words $A \oplus B = 1 \iff A \neq B$, and $A \oplus B = 0$ otherwise.

We can now define our statistical tests, namely the frequency and autocorrelation tests.

Definition 6.2 (Frequency Test). For a binary sequence $S = s_1,..., s_n$ where $s_i \in \{0,1\}$, let $X = \sum_{i=1}^{n} s_i$ be the number of 1's. Under the null hypothesis of randomness, each s_i is 0 or 1 with probability 1/2, so $X \sim$ Binomial(n, 1/2). Define O_k to be the observed frequency of some outcome $k \in \{0,1\}$, with $O_1 = X$ (number of 1's) and $O_0 = n - X$ (number of 0's). The expected frequency E_k is $E_0 = E_1 = n/2$, as 0's

and 1's are equally likely. The chi-square statistic measures deviation from expected randomness:

$$\chi^2 = \sum_{k=0,1} \frac{(O_k - E_k)^2}{E_k}.$$

By construction, the statistic above follows a chi-square distribution that has a single degree of freedom. The p-value, which can be defined as the probability of observing a χ^2 value that is at least as extreme under the null hypothesis, is then calculated using the chi-square distribution's cumulative distribution function. A p-value that is below an agreed-upon threshold, usually $p = 0.01$, means that we can reject the null hypothesis and conclude that there is non-randomness.

Definition 6.3 (Autocorrelation Test). For a binary sequence $S = s_1,..., s_n$ where $s_i \in \{0,1\}$, define a *shift* d (where $1 \leq d < n$) to be an operation that offsets the sequence by d positions, thus allowing us to compare s_i with s_{i+d}. The autocorrelation coefficient for shift d is defined to be:

$$A(d) = \sum_{i=1}^{n-d} (s_i \oplus s_{i+d}),$$

where $\oplus$ is defined as the XOR operation in definition 6.1. Then by definition 6.1, $s_i \oplus s_{i+d} = 1$ if $s_i \neq s_{i+d}$ and 0 otherwise. From this, we infer that $A(d)$ allows us to calculate the number of positions in which the original binary sequence and the shifted sequence differ. Under our null hypothesis (of randomness), each computation that yields $s_i \oplus s_{i+d} = 1$ has a probability of 1/2, thus we can say that $A(d) \sim$ Binomial(n - d, 1/2), with expected value $\frac{(n-d)}{2}$ and variance $\frac{(n-d)}{4}$. We then define the test statistic ζ to be as follows:

$$\zeta = \frac{|A(d) - (n-d)/2|}{\sqrt{(n-d)/4}}.$$

Observe that ζ measures the deviation of $A(d)$ from its expected value(normalized by the standard deviation). For sufficiently large n - d, this approximates to a standard normal distribution. Then if ζ exceeds a threshold under such a distribution, we know that there exists a significant correlation between bits separated by d, and this indicates non-randomness.

Such tests are necessary to test for randomness alongside techniques such as the Chernoff bounds. They are quite handy when deciding if a given ciphertext is likely to leak patterns in plaintext. It is precisely such tests that allows us to determine to what degree an encryption scheme is deterministic (Menezes et al., 1997; Rukhin et al., 2010).

Applications

The probabilistic techniques in this paper are crucial to many cryptographic systems, and we shall illustrate a few such applications. Consider entropy, as discussed in Section 2. One of the most impressive applications of entropy within cryptography is its role in generating keys for the Advanced Encryption Standard (AES), which is the symmetric-key block cipher used by the U.S. National Institute of Standards and Technology (NIST). AES operates on 128-bit blocks, and it allows for key lengths of 128, 192, and 256 bits. This allows for seeds with relatively high levels of entropy, which allow them to resist brute force attacks in protocols that govern internet communication, such as TLS (Transport Layer Security) (Menezes et al., 1997). For example, 256-bit keys with $H(K) = \log(2^{256}) \approx 177.4$ nats, are ideal. Now consider Chernoff bounds, as discussed in Section 4. Chernoff bounds allow for offensive cryptographic ability. For example, they are used in side channel attacks, as briefly described in Section 4. Such attacks use Chernoff bounds to detect non-random power consumption patterns in people's smartcards, and subsequently reveal key bits with high levels of confidence. They do this by measuring deviations in power consumption or the timing required to infer key bits, thus allowing negligible failure probabilities in cryptographic protocols. Thus, by giving us methods and techniques to measure randomness, probability theory has massively furthered the field of cryptology.

References

Blum, L., Blum, M., & Shub, M. (1986). *A Simple Unpredictable Pseudo-Random Number Generator. SIAM Journal on Computing*, *15*(2), 364–383.

Burroughs, J., Lieberman, D., & Reeds, J. (2009). *The Secret Life of Andrew Gleason. Notices of the American Mathematical Society*, *56*(10), 1239–1243.

Chernoff, H. (1952). *A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations. Annals of Mathematical Statistics*, *23*(4), 493–507.

Goemans, M. X. (2015). *Chernoff Bounds, and Some Applications.* Lecture Notes, MIT 18.310A. Available at: https://math.mit.edu/~goemans/18310S15/chernoff-notes.pdf.

Goldwasser, S., & Micali, S. (1984). *Probabilistic Encryption. Journal of Computer and System Sciences*, *28*(2), 270–299.

Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1997). *Handbook of Applied Cryptography.* CRC Press.

Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., & Vo, S. (2010). *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications.* NIST Special Publication 800-22, Revision 1a.

Shannon, C. E. (1948). *A Mathematical Theory of Communication. Bell System Technical Journal*, *27*(3), 379–423.

Shannon, C. E. (1949). *Communication Theory of Secrecy Systems. Bell System Technical Journal*, *28*(4), 656–715.

Bioremediation of Effluents: Developing a Bioadsorbent Medium to Treat Water Heavy Metal and Textile Dye Pollutants Based on the Combined Adsorption by Neem and Other Biosorbents

Adnan Bin Alamgir and Golam Yasin
Harvard College '29 and Dhaka Residential Model College '25

With the expansion of textile industries and their extensive use of chemicals, industrial wastewater contaminated with heavy metals and textile dyes is finding its way into our usable water sources more frequently than ever. The high cost of current treatment procedures, as well as the environmental concerns associated, puts our drinkable water at risk of contamination. Agricultural byproducts, or biosorbents, can function as bio-adsorbents to remove dissolved substances from wastewater through adsorption. These materials are abundant and organic but are usually underutilized due to their generated sludge and lack of extensive efficiency. The purpose of our research is to develop a contact adsorbent medium by processing and combining five well-known biosorbents: neem (*Azadirachta indica*) leaf and bark, peanut (*Arachis hypogaea*) husk, coffee (*Coffea arabica*) ground, watermelon (*Citrullus lanatus*) rind, and banana (*Musa sapientum*) peels. By combining their treatment capacities, we were able to treat copper-contaminated water. We found that a paper-shaped medium can efficiently remove a significant portion of dissolved material while minimizing sludge, thus creating potential for replacing our existing non-sustainable treatment materials with an organic alternative derived from agricultural wastes.

Introduction

The influx of heavy metal substances into the environment has been increasing at a concerning rate over the past few centuries, with the majority of the freshwater bodies near cities and human habitats passing the safe levels set by World Health Organization (Panchamoorthy et al., 2024; WHO, 2017). Almost a third of the world's population still lacks clean and safe drinking water, one major cause of which is heavy metal pollution. In Bangladesh, an estimated 70 million people who consume groundwater are exposed to unsafe levels of arsenic (WHO, 2017). In addition to heavy metals, production of over a million tons of dyes and pigments takes place annually worldwide. The effluents from dyeing industries are contaminated with highly colored compounds (e.g., azo dyes such as malachite green, aniline yellow, and methylene blue) and have a high chemical and biological oxygen demand (Garg et. al., 2004). If allowed into usable water bodies, they can harm aquatic ecosystems and human beings alike due to their carcinogenic, allergenic, and toxic effects.

The use of adsorbent organic materials in treating effluents is a well-studied phenomenon, with many works of research focused on a particular pair of adsorbent and adsorbates. In Garg et al. (2004), rosewood sawdust pretreated with sulfuric acid is used to treat methylene blue, a common blue dye in textile industries. Sawdust is a very common byproduct of the timber industry, so their approach also fulfilled the conditions of low cost and sustainability. Khattri et al. (2009) take on a similar approach using neem sawdust and targeting malachite green dye. While some of these studies focus on very specific topics, studies such as the one done by Farias et al. (2023) look at the overall porous adsorption capacity of a specific biosorbent, namely banana peel powder. Most of the existing literature on organic bioadsorbents focuses on the specific adsorption capacities of a given biosorbent, which, due to the physical and chemical properties of that specific biosorbent, act primarily well on a particular dye or heavy metal component. By combining these individual biosorbents in a manner that preserves their individual adsorbent properties, it is possible to create a more versatile effluent-treating medium.

In this project, we combined a range of organic materials—neem leaf powder (NLP) and sawdust, peanut shell powder (PSP), banana peel powder (BPP), coffee grounds (CG), and watermelon rind (WMR)—to remove heavy metals and common textile dyes through adsorption. Adsorption is the physicochemical process in which molecules from an aqueous solution (adsorbates) adhere to the adsorbent surface. The primary ingredient, neem, contains high levels of flavonoids and tannins, which assist in the effective removal of heavy metals from water (Khattri et al., 2009). PSP has a structural formation consisting mainly of cellulose networks, making it quite effective in the removal of turbidity (Nkansah et al., 2019). BPP is known for decontaminating wastewater, mainly in the removal of textile dyes, heavy metals, and even crude oil (Farias et. al., 2023). Another common waste material, CG, has a lignocellulosic composition of 51.5% cellulose and 40.6% hemicellulose, which makes it an efficient bioadsorbent (Mukti & Hidayat, 2019). WMR is composed of carbonaceous materials (e.g., cellulose, pectin, etc.), making it a suitable adsorption agent for metals and dyes (Bhattacharjee et al., 2020). The primary causes of physical adsorption are the van der Waals and electrostatic forces between the adsorbate molecules and the adsorbent surface. Using a paper-like structure maximizes the surface area while keeping the composition to a minimum. It also creates a reinforced structure that can work without breaking down into suspension as soon as it encounters water, which, along with an extra biopolymer layer covering the paper, significantly minimizes sludge.

Each of the discussed bioadsorbents can work as effective removal agents for pollutants, and in this paper, their individual remedial potencies were combined into a paper-shaped contact adsorption medium to create an efficient solution for wastewater remediation that performs on the same level as existing adsorbent media, with over a 50% reduction in treatment costs and reduced environmental concerns. Furthermore, the adsorption capacity of this medium was analyzed in the theoretical framework of adsorption isotherms, which is an empirical model of an equilibrium function between adsorbent concentration and the amount of adsorption at a given temperature. By fitting our limited data points on these empirical models, we can predict the behavior of the adsorption mechanism. The Langmuir isotherm is one such model that well describes adsorption in a solid/liquid interface, an interface similar to our case, and the bioadsorbent paper indeed is found to fit well to this model.

Methods

In designing the paper to acquire maximum adsorption with minimum dissociation, we chose a simplistic design where the bioadsorbent paper has two individual layers—one adsorbent layer and one protecting layer (Fig. 1). The outer layer is made of a biopolymer that creates a thin separation between the bioadsorbent paper and the solution. As adsorption does not require chemical interaction, the dye molecules and metal ions attracted by the bioadsorbent layer accumulate on the polymer, and the inner layer stays intact and reusable. We do expect the paper to degrade and will eventually need to be replaced, but because of the polymer layer, this will be substantially delayed.

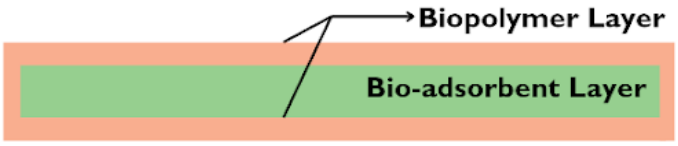


Figure 1. Schematic diagram of the bioadsorbent paper. The inner layer is composed of a mixture of plant byproducts from *Azadirachta indica*, *Arachis hypogaea*, *Coffea arabica*, *Musa sapientum*, and *Citrallus lanatus*. The outer layer is a biopolymer made from agar and glycerin.

Material Processing and Papermaking

The initial part of our project was preparing the bioadsorbent paper, for which we had to first process the materials. For neem, we washed the leaves with distilled water and dried them, first leaving them outside for two days and then in an oven at 100°C for an hour. After that, we blended the leaves and strained the powder until we were left with fine-grained NLP (Nkansah et al., 2019). For neem sawdust, PSP, BPP, and WMR, we carried out similar processes of drying, blending, and straining. For CG, we only had to dry them to use for our purposes. For the papermaking step, we mixed the processed materials in the desired ratio by mass (Table 1) and blended them together with some water to make a pulp. In deciding on the effective ratio, the two factors we considered were how available the biosorbents were locally (year-round agricultural products compared to seasonal products) and the structural integrity. For example, WMR is not only a seasonal agricultural waste but also more prone to liquefying the pulp too much to have the proper density to become a paper-like structure. We added okra extract as a binding agent to keep the paper from breaking down, courtesy of its effective glue-like property. We ran the mixture through a tub of water and into a paper mold. After letting it dry for a day, we carefully separated the sheet of bioadsorbent paper. For the outer layer, we mixed agar and water to work as a substitute for starch and then processed it with glycerin to create a polythene resembling a biopolymer.

Biosorbent	Neem	PSP	BPP	Coffee ground	WMR
%[a]	40	20	20	10	10

Table 1. Percent by mass composition of biosorbents in the bio-adsorbent paper.

Preparation of the Removal Rate Test Apparatus

Our efficiency testing was done in two steps: NLP/neem sawdust testing and bioadsorbent paper testing. For testing, we used a controlled molarity solution of the heavy metal salt copper (II) sulfate ($CuSO_4.5H_2O$). In the first test, using a weight balance with an accuracy of a thousandth of a gram, we measured 2.00g of $CuSO_4$ salt and made five beakers of 150mL solution, giving the total volume of 150mL × 5 = 750mL. Given that the molar mass of $CuSO_4.5H_2O$ is 249.5, the molarity, M, of our solution was,

$$\mathrm{M} = \frac{2\ \mathrm{g}}{249.5\ \mathrm{g/mol}} \times 0.750\ \mathrm{L} = 0.011\ \mathrm{M} \quad (1).$$

Beaker 1 was kept as the reference beaker; in beaker 2 and 3, we weighed and added 1.800g of NLP; in beaker 4 and 5, we added 1.200g of neem sawdust. We added 1 mL of dilute hydrochloric acid (HCl) to work as a pH activator (Kalam et al., 2021). After leaving the beakers for two hours, we filtered out the precipitated substance using a filter paper for further analysis. For testing the bioadsorbent paper, we used an isomolar solution of copper sulfate, adding 50 cm^2 (5 cm x 10 cm) pieces of the paper, weighing on average 2.460 g.

Removal Rate Testing

The primary and crudest method of our testing was to observe the change in the solution's color. Using copper (II) sulfate for testing has an advantage due to its blue hue, whose strength is approximately proportional to Cu^{2+} concentration. The next test was weight comparison; comparing the final precipitate (or paper) weight with the pre-adsorption weight gives an estimate of the removal rates. The differences were precisely measurable as we used a high-concentration solution. Even though real-life wastewater will likely have much lower levels of copper concentration, having a high molarity helps to reduce the uncertainty in our tests while preserving the relative removal capacities. To measure the removal capacity for lower concentrations, a more sophisticated testing apparatus is required (see discussions). In each of the beakers, there were 0.4g of $CuSO_4.5H_2O$. Therefore the weight of Cu^{2+}, w_{Cu}, was

$$w_{Cu} = \frac{0.4\ \mathrm{g} \times 63.5\ \mathrm{g/mol}}{249.5\ \mathrm{g/mol}} = 0.102\ \mathrm{g} \quad (2).$$

We then calculated Cu^{2+} removal using the following equation, where w_i is the initial mass and w_f is the final mass,

$$Cu^{2+}\ \mathrm{removal} = \frac{w_f - w_i}{w_{Cu}} \times 100\% \quad (3).$$

Lastly, we did an ion test with potassium ferrocyanide ($K_4[Fe(CN)_6]$), which reacts with Cu^{2+} to produce copper (II) hexacyanoferrate ($Cu_2[Fe(CN)_6]$), a reddish brown precipitate.

$$CuSO_4 + K_4[Fe(CN)_6] \rightarrow Cu_2[Fe(CN)_6] + K_2SO_4 \quad (4).$$

We took 4 mL of solution from the beakers into a batch of test tubes and added 1 mL of $K_4[Fe(CN)_6]$. The level of precipitation gives an effective estimate of metal ion removal.

Ion Identification Test

There was a significant visual difference in $Cu_2[Fe(CN)_6]$ precipitation in the two test tubes (Fig. 2b), as well as in reaction rate. In the treated solution, the reaction took place rapidly, but in the reference solution, it took a much longer time, indicating a notable difference in Cu^{2+} concentration.

Weight Comparison Test

For this test, we first carry out the test for unprocessed (not processed into the bio-adsorption paper) neem sawdust, which is the largest constituent element of our bioadsorbent paper and can give us a reference to measure the efficacy of giving the biosorbent a paper-like structure.

NLP/Sawdust Testing: After measuring the weights for beakers 2-5, we used equation (3) to calculate the removal rates. Table 2 shows the results.

Adsorbent	Initial Weight (g)	Final Weight (g)	Cu^{2+} Removal (%)	Average Removal Rate (%)
NLP	1.800	1.863	61.7	64.2
NLP	1.800	1.868	66.7	
Sawdust	1.200	1.274	72.5	71.6
Sawdust	1.200	1.272	70.6	

Table 2. Measured removal rates for neem leaf powder (NLP) and neem sawdust.

Bioadsorbent Paper Testing: We weighed the paper before and after adsorbent uptake, and conducted three tests for a statistically significant result. The removal rates in each test and the average rates are presented in Table 3. In both cases, the reference beakers (beaker 1) were used to make sure there are no implicit factors other than the bioadsorbent contributing to the heavy metal removal, and indeed, no precipitations were observed in the reference beakers.

Labels	Paper Area (cm^2)	Initial Weight (g)	Final Weight (g)	Cu^{2+} Removal (%)	Average Removal Rate (%)
Pap-1	50	2.443	2.520	75.5	76.5
Pap-2	50	2.426	2.510	82.4	
Pap-3	50	2.514	2.587	71.6	

Table 3. Measured removal rate for bio-adsorbent paper.

Finally, we repeated the test for a continuous set of removal rates with time, to fit a graph for the adsorption mechanism under constant temperature. Paper sizes of 50 cm^2 and 40 cm^2 were used for two different graphs. Using the data points, we fitted the best fit curves, which are shown in Fig. 3. The isotherms showed a similar trend to Langmuir isotherms, which was reasonable given the monolayer and homogeneous surface of the bioadsorbent paper (Kalam et al., 2021).

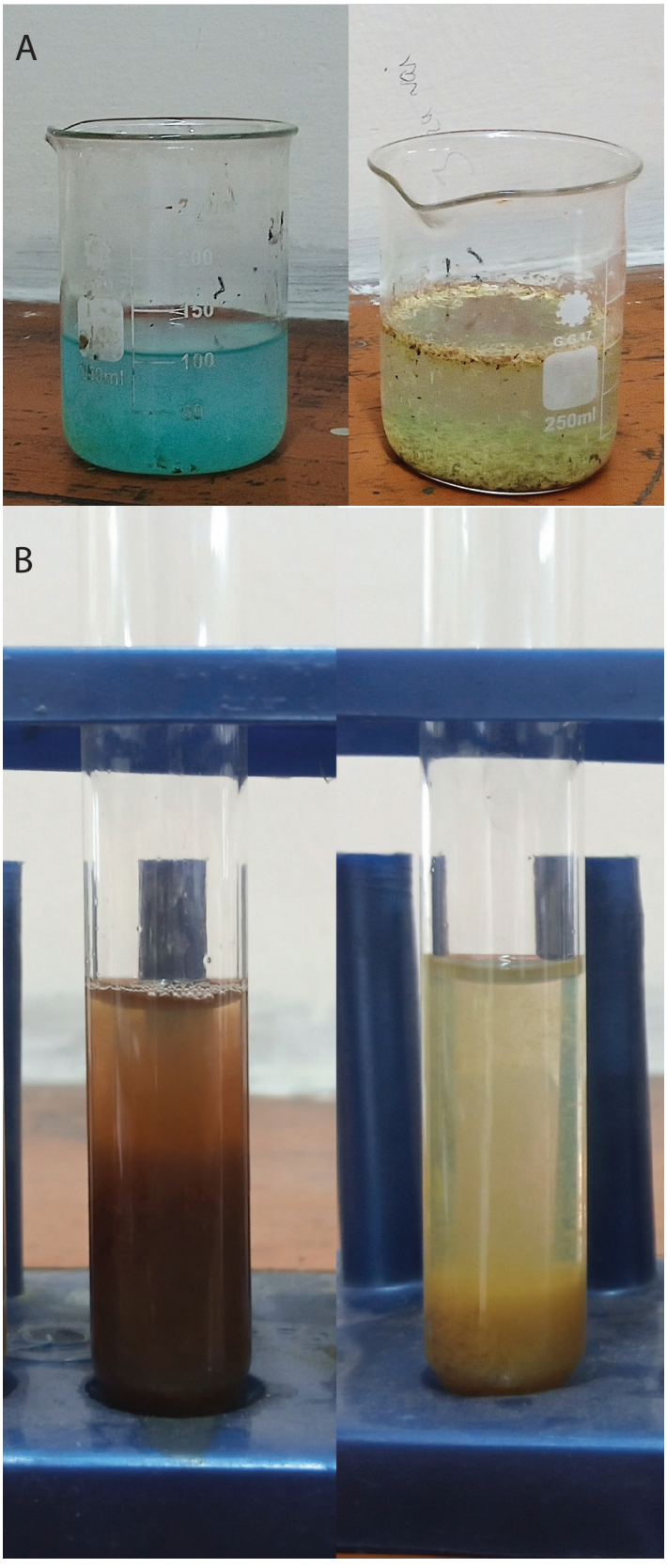


Figure 2. (a) Visual results from the color test. (b) Visual results from the ion test. In both images, the left and right parts contain reference and treated solutions, respectively.

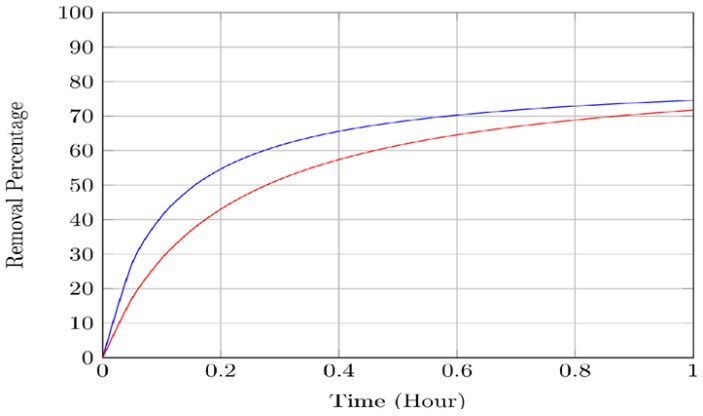


Figure 3. Adsorption mechanism isotherms for 40 cm^2 (red) and 50 cm^2 (blue) sized pieces of the bioadsorbent paper.

Discussion

The main objective of our project was to develop an adsorbent medium by combining a wide range of biosorbents capable of water remediation. Our bioremediation technique has shown the ability to remove heavy metals from water at around 80% efficiency without creating any substantial sludge. Although not explicitly tested, we expect this adsorption medium to work well in removing textile dyes as well, given the ionic constitution of the majority of dyes (McCord & Roza, 2018). Our experiments showed higher removal rates in slightly acidic solutions, and NLP and sawdust have been found to work best in the pH range of 4 to 6 (Kalam et al., 2021). The reason we focused more on an acidic apparatus is due to the basic nature of dissolved metallic ions, as well as the fact that most textile dyes tend to be basic. Carrying out the experiments in a neutral or basic solution instead, while it may lower the adsorption capacity, will give us more information on how removal rates and pH values may be connected. A number of past studies also look at this correlation, including Garg et al. (2004) and Febriana et al. (2009), and although their pH ranges do not necessarily always overlap, they find better removal rates in an acidic environment. More acidic is not necessarily better, as the extra protons (H^+) present in the solution start to compete with Cu^{2+} ions for the active sites. Temperature is another factor; as adsorption is an exothermic process ($\Delta H < 0$), the removal rates should slightly increase with increasing temperature (Kalam et al., 2021). Surface properties such as surface area, polarity, and pore size distribution can also influence adsorption rates. Designing to maximize these properties should facilitate the removal capacity. In our work, we have used a high-molarity test solution for a higher level of accuracy. Real-life contaminated water usually has a much lower concentration of contaminants, in the mg/L levels, compared to the g/L levels tested in this paper, and methods such as atomic absorption spectrometry (AAS) can further validate our results.

Unlike chemical treatment methods, the discussed biosorbents do not introduce harmful chemicals into our waters, making them safe for drinking and irrigation. These biosorbents can be produced locally almost everywhere in the world, reducing the overall cost and the environmental impact associated with transportation. As these materials are mostly discarded as agricultural waste, utilizing them helps waste management and promotes recycling. Moreover, the process is similar to industrial papermaking and can be mass-produced with minimum effort. The usability of our bioadsorbent paper is versatile, on both the industrial and personal scale. Regular-sized strips could be used to treat water to remove the bulk of the heavy metals and textile dyes from water in the event of flooding or an industrial spill. For industries such as textiles, this paper can be incorporated in the tertiary stage of the conventional effluent treatment process.

In this paper, we have discussed the design, constituents, and implementation of an adsorption medium made entirely from organic biosorbents. This can provide a pragmatic solution to the high expense of industrial wastewater treatment, due to which many middle- to small-scale industries fail to treat their effluents. Using our bioadsorbent paper can reduce environmental harm and waterbody degradation, as well as providing a cost-effective solution for industries to treat their effluents and help create a sustainable future.

Acknowledgments

We would like to thank the Department of Chemistry at Dhaka Residential Model College for providing the chemical equipment necessary for our experiments.

References

Bhattacharjee, Chiranjit, Suman Dutta, and Vinod K. Saxena. "A Review on Biosorptive Removal of Dyes and Heavy Metals from Wastewater Using Watermelon Rind as Biosorbent." *Environmental Advances 2*, (2020): 100007. Accessed August 30, 2024.

Farias, Kelly C S et al. "Banana Peel Powder Biosorbent for Removal of Hazardous Organic Pollutants from Wastewater." *Toxics* vol. 11,8 664. 1 Aug. 2023.

Febriana, Novie, et al. "Neem Leaf Utilization for Copper Ions Removal from Aqueous Solution." *Journal of the Taiwan Institute of Chemical Engineers*, vol. 41, no. 1, 2009, pp. 111-114.

Garg, V.K., et al. "Basic Dye (Methylene Blue) Removal from Simulated Wastewater by Adsorption Using Indian Rosewood Sawdust: A Timber Industry Waste." *Dyes and Pigments*, vol. 63, no. 3, 2004, pp. 243-250.

Kalam, Shams et al. "Surfactant Adsorption Isotherms: A Review." *ACS Omega. XXXX.* (2021). 10.1021/acsomega.1c04661.

Khattri, S.D., and M.K. Singh. "Removal of Malachite Green from Dye Wastewater Using Neem Sawdust by Adsorption." *Journal of Hazardous Materials 167*, no. 1-3 (2009): 1089-1094.

Mukti, Nur Indah and Hidayat, Arif. (2019). "Characterization of coffee grounds as biosorbent for removal of dyes from aquaeous solutions." IOP Conference Series: Materials Science and Engineering. 625. 012031.

Nkansah, Marian, Moses Donkoh, Osei Akoto, and James Ephraim. (2019). "Preliminary Studies on the Use of Sawdust and Peanut Shell Powder as Adsorbents for Phosphorus Removal from Water." Emerging Science Journal. 3. 33.

Panchamoorthy Saravanan et al., "Comprehensive review on toxic heavy metals in the aquatic system: sources, identification, treatment strategies, and health risk assessment", *Environmental Research*, Volume 258, 2024, 119440, ISSN 0013-9351, 10.1016/j.envres.2024.119440.

World Health Organization (2017). *Progress on drinking-water sanitation and hygiene*. WHO, Geneva, Switzerland.

Yazdani McCord, Maryam Roza. (2018). Engineered adsorptive materials for water remediation - Development, characterization, and application.

Shifted Symmetric Functions and Young Tableaux

Zak Adams and Victor Seco Roopnaraine
Harvard College '26

We provide an expository overview of shifted symmetric functions by focusing on their evaluation on Young diagrams and their expression in multirectangular coordinates. Shifted Schur functions are shown to be polynomials in these coordinates, with nonnegative coefficients when expanded in the falling factorial basis. Motivated by classical results on symmetric functions, we then explore relations and generalizations to Jack polynomials. This paper surveys the work of Alexandersson and Féray (2017) and the frameworks provided by Stanley (2001, 2002) and Okounkov–Olshanski (1997a, 1997b).

Introduction

1.1 Motivation

Shifted symmetric functions arise as deformations of classical symmetric functions, where symmetry is considered with respect to the shifted variables. Their study, initiated by Okounkov and Olshanski (1997b), revealed connections with structures from classical symmetric function theory, including Schur functions, Kostka numbers, and character formulas.

While classical symmetric functions are invariant under arbitrary permutations of variables, shifted symmetric functions are formal power series symmetric in the shifted variables. Such functions can be evaluated on Young diagrams, which establish a relation between algebraic combinatorics and the representation theory of symmetric groups.

Among shifted symmetric functions, the shifted Schur functions $S^\star_\mu$ and the functions K_μ related to shifted Kostka numbers are of interest. Both are determined by their evaluations on Young diagrams in multirectangular coordinates and have nonnegative coefficients when expressed in terms of falling factorials.

We introduce multirectangular coordinates through the work of Stanley (2001, 2002). He describes a Young diagram via two sequences of non-negative integers, $(p_1, \cdots, p_d)$ and $(r_1, \cdots, r_d)$. Expressing shifted symmetric functions in such coordinates gives positivity.

The paper aims to provide an expository overview of the theory of shifted symmetric functions and to study their polynomial structure in multirectangular coordinates. We also discuss its relations to Jack polynomials (Goulden & Jackson, 1996; Lassalle, 2009).

1.2 Format

In Section 2, we define shifted symmetric functions and their connection with Young diagrams, allowing us to introduce the shifted Schur functions $S^\star_\mu$, following the constructions of (Okounkov & Olshanski, 1997a). In Section 3, we follow Stanley's work (2002) in defining set partitions and multirectangular coordinates in order to prove the shifted Schur function has nonnegative rational coefficients in the falling factorial basis (Alexandersson & Féray, 2017, Theorem 1.2). Section 4 applies a similar proof but to the shifted Jack polynomial (Alexandersson & Féray, 2017, Theorem 1.5).

1.3 Main Results

We summarize here the main results of this paper. The precise definitions of notation will be provided in later sections.

- Theorem 3.1 (Okounkov & Olshanski, 1997b; Stanley, 2001): If F is an α-shifted function, as in Definition 2.1, then when evaluated on a Young diagram λ, its value F_λ, expressed in multirectangular coordinates (p_i, r_i)(see Definition 3.5) is a polynomial in (p_i, r_i) with coefficients in $\mathbb{Q}(\alpha)$.

1	2	4	7
3	5	8	
6			

Figure 1. An example of a standard Young tableau of shape $\lambda = (4,3,1)$.

- Theorem 3.2 (Alexandersson & Féray, 2017, Theorem 1.2): For each partition μ, the shifted Schur function, $S^\star_\mu\left(\mathbf{r}^p\right)$, as defined in Definition 2.2, admits an expansion with non-negative rational coefficients in the falling factorial basis $\mathfrak{B}$ (see Definition 4.2).
- Theorem 4.1 (Alexandersson & Féray, 2017, Theorem 1.5): For every integer k, the shifted Jack polynomial, denoted $J^{\star,(\alpha)}_{(k)}\left(\mathbf{r}^p\right)$, given in Definition 4.1, has non-negative rational coefficients when expressed in the α-falling factorial basis.

Shifted Symmetric Functions and Young Tableaux

2.1 Shifted Symmetric Functions

Definition 2.1 (Shifted Symmetric Function) (Alexandersson & Féray, 2017).An α-shifted symmetric function F is a sequence $(F_N)_{N\geq 1}$ such that

1. For each $N \geq 1$, is F_n a polynomial in N variables $(x_1, \cdots, x_N)$ with coefficients in $\mathbb{Q}(\alpha)$ that is symmetric in $x_1 - \frac{1}{\alpha}, x_2 - \frac{2}{\alpha}, \cdots, x_N - \frac{N}{\alpha}$.
2. For each $N \geq 1$, $F_{N+1}(x_1, \cdots, x_N, 0) = F_N(x_1, \cdots, x_N)$ (this is known as the stability property.)
3. $\sup_{N\geq 1} \deg\left(F_N\right) < \infty$.

Definition 2.2 (Shifted Schur Function). For any $\mathrm{Par} \ni \mu$ and $n \geq 1$, we define the shifted Schur function $S^\star_\mu$ in variables $(x_1, \cdots, x_n)$ by:

$$S^\star_\mu(x_1, \ldots, x_n) = \frac{\det\left((x_i + n - 1)_{\mu_j + n - j}\right)}{\det\left((x - i + n - i)_{n-j}\right)}.$$

We note that the highest degree terms of $S^\star_\mu$ is exactly the Schur function S_μ. In addition to Subsection 2.3, we can observe that well-known formulas involving Schur functions naturally extend themselves to shifted Schur functions. Such formulas include Jacobi-Trudi identity and the combinatorial expansion in terms of semi-Standard Young tableaux (Okounkov & Olshanski, 1997a).

The above serves as an example of a 1-shifted symmetric function.

2.2 Young Tableaux

Define a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ to be a finite, weakly decreasing sequence of non-negative integers. The associated Young diagram is a left-justified array of boxes, with λ_i boxes in the *i*-th row from the top. We identify a partition with its Young diagram.

Definition 2.3 (Standard Young Tableau). A standard Young tableau of shape λ is a filling of the Young diagram of λ with numbers $1, 2, \ldots, |\lambda|$ such that entries increase along each row (left to right) and increase down each column (top to bottom). For example, the following is a standard Young tableau of shape $\lambda = (4, 3, 1)$.

Definition 2.4 (Semi-standard Young Tableau). A semi-standard Young tableau (SSYT) of shape λ is a filling of the Young diagram with positive integers such that entries weakly increase along each row and strictly increase down each column. We define the type of an SSYT to be the weak composition $\tau = (\tau_1, \tau_2, \ldots)$ where τ_i counts the number of boxes filled with the integer *i*.

We then have that the Kostka number K^λ_τ counts the number of SSYTs of shape λ and type τ (Macdonald, 1995).

2.3 Shifted Symmetric Functions and Tableaux

The connection between shifted symmetric functions and Young tableaux yields results that do not occur in the non-shifted setting. In fact, the evaluation $S^\star_\mu(\lambda)$ of a shifted Schur function equals 0 unless λ contains μ. When λ contains μ, we find $S^\star_\mu(\lambda)$ counts, up to normalization, the number of standard tableaux of skew shape λ/μ (Okounkov & Olshanski, 1997a).

As $S^\star_\mu(\lambda)$ and $K_\mu(\lambda)$ have tableau interpretations, we can understand shifted symmetric functions as polynomial versions of tableau counting. Moreover, if we express Young diagrams using multirectangular coordinates, we can show that shifted symmetric functions become polynomials in multirectangular coordinates with positivity properties.

Multirectangular Coordinates and Shifted Schur Functions

3.1 Set Partitions

We refer to Section 3.1 of (Alexandersson & Féray, 2017) for the definitions of set partitions.

Definition 3.1 (Set Partition). A set partition of a finite set E is a collection of non-empty, pairwise disjoint subsets $S = \{S_1, S_2, \ldots, S_s\}$ whose union is E. Each subset S_i is called a block and the number of blocks is given by $s = |S|$.

The order of set partitions is defined by the *refinement order*. Given two partitions of E denoted $\widetilde{S}$ and S, we say $\widetilde{S} \leq S$ if every block of $\widetilde{S}$ is contained entirely within some block of S. In this case, we write $\widetilde{S}$ as finer than S, and S is coarser than $\widetilde{S}$.

Remark 3.1. The set of all partitions of E, ordered by refinement, forms a lattice. In this lattice, any two partitions have a greatest lower bound (their common refinement) and a least upper bound (their common coarsening).

Definition 3.2 (Induced Partition). Let $\widetilde{S} \leq S$ be partitions of E, and let S_i be a block of S. The induced partition $S/\widetilde{S}_i$ is the set partition formed by the blocks of $\widetilde{S}$ that are contained in S_i.

Intuitively, we may see that within each block S_i of S, the partition $\widetilde{S}$ may split S_i into smaller pieces. Such pieces form a partition of S_i.

Proposition 3.1. Given a partition S of E, if for each block S_i we choose a partition $\widetilde{S}_i$ of S_i, then the union of the partitions $\widetilde{S}_i$ forms a partition $\widetilde{S}$ of E that is finer than S.
Conversely, given a partition $T = \{T_1, \cdots, T_t\}$ of the blocks of S, we can form a coarser partition of E by taking unions,

$$T_j \mapsto \bigcup_{S_i \in T_j} S_i.$$

Thus, partitions of the set of blocks correspond naturally to partitions of E that are coarser than S.

We now introduce how functions relate to set partitions.

Definition 3.3 (Functions and Induced Partitions). Let S be a partition of E and let $f : S \to D$ be a function assigning a value in D to each block of S. This induces a function $\overline{f} : E \to D$ by setting $\overline{f}(e) = f(S_i)$ if $e \in S_i$.

Definition 3.4 (Partitions from Functions). Conversely, given any function $g : E \to D$, we define the associated set partition

$$\Pi(g) = \{g^{-1}(\{x\}) : x \in g(E)\},$$

where two elements e, e' are in the same block of $\Pi(g)$ if and only if $g(e) = g(e')$.

Proposition 3.2. Let S be a set partition of E and let $f : S \to D$ be a function. Then:

$$S \leq \Pi\left(\overline{f}\right),$$

meaning that S is finer than the partition induced by $\overline{f}$.

3.2 Multirectangular Coordinates

Multirectangular coordinates were introduced by Stanley (2001, 2002) and provide a natural way to parametrize Young diagrams. This coordinate system allows us to express evaluations as explicit polynomials in combinatorial parameters associated with the shape of the tableau. This pays dividends when working with shifted symmetrical functions.

Definition 3.5 (Multirectangular Coordinates). Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ be a partition. The multirectangular coordinates of λ consists of two sequences of non-negative integers

$$(p_1, \ldots, p_d), \quad (r_1, \ldots, r_d)$$

satisfying the following:

- $p_i > 0$ for all i;
- $r_1 \geq r_2 \geq \ldots \geq r_d \geq 0$;
- The rows of λ are grouped into d blocks with

$$\lambda = \left(\underbrace{r_1 + r_2 + \ldots + r_d}_{p_1 \text{ times}}, \underbrace{r_2 + r_3 + \ldots + r_d}_{p_2 \text{ times}}, \ldots, \underbrace{r_d}_{p_d \text{ times}} \right).$$

If we require each block to group together consecutive rows of equal length and that each grouping is maximal (i.e., no two consecutive blocks can be merged without violating the decreasing width condition, then our multirectangular coordinates are uniquely determined.

Example 3.1. Take $\lambda = (5, 3, 3, 2)$. The first row has 5 boxes, the second and third row each has 3 boxes, and the fourth row has 2 boxes. This gives the multirectangular decomposition of $(p_1, p_2, p_3) = (1, 2, 1)$ and $(r_1, r_2, r_3) = (2, 1, 2)$. Graphically, the Young diagram can be visualized as the combination of three rectangles.

Let us now consider several properties of multirectangular coordinates.

Lemma 3.1 (Uniqueness). The multirectangular decomposition of a Young tableau under the maximal grouping convention is unique.

Proof. Given a Young diagram, we proceed bottom-up. Start with the bottommost rows, identifying the maximal set of consecutive rows having the same length. Then define p_d as the number of these rows and r_d to be their common row length. Remove these rows and repeat the process recursively on the remaining upper diagram. As Young diagrams are weakly decreasing in row length, and as each step groups maximally, no ambiguity arises in such a construction. □

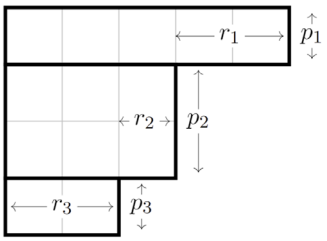


Figure 2. Multirectangular decomposition of $\lambda = (5,3,3,2)$ after reversing the rectangle order.

3.3 Polynomiality in Multirectangular Coordinates

An important result concerns the evaluations of shifted symmetric functions. They become polynomials in the multirectangular coordinates.

Theorem 3.1 (Okounkov & Olshanski, 1997a; Stanley, 2001). Let F be an α-shifted function as in Definition 2.1. When evaluated at a Young diagram λ, expressed via multirectangular coordinates (p_i, r_i), the value F_λ lies in $\mathbb{Q}(\alpha)[p_1, \ldots, p_d, r_1, \ldots, r_d]$.

Proof. By definition, each F_N is a polynomial in $\{x_1, \ldots, x_N\}$, symmetric in the shifted variables

$$x_1 - \frac{1}{\alpha}, \quad x_2 - \frac{2}{\alpha}, \quad \ldots, \quad x_N - \frac{N}{\alpha}.$$

When we evaluate F at a partition λ, we substitute $(x_1, x_2, \cdots) = (\lambda_1, \lambda_2, \cdots)$, padding with zeros if necessary. In multirectangular coordinates, the part $\lambda_i$corresponding to a row in block k can be expressed as

$$\lambda_i = r_k + r_{k+1} + \cdots + r_d.$$

Moreover, the index i (the row number) can be written explicitly in terms of the cumulative sums of the p_j. Hence, each shifted variable $\lambda_i - \frac{i}{\alpha}$ is an affine function of the p_j and r_j, with rational coefficients depending on α. Since polynomials are closed under affine substitution, and the stability property ensures compatibility across $\mathbb{N}$, it follows that $F(\lambda)$ is a polynomial in (p_i, r_i) with coefficients in $\mathbb{Q}(\alpha)$. □

Finally, let us emphasize that multirectangular coordinates provide a combinatorial basis for describing shifted functions. This comes from each block p_i corresponding to a group of rows of constant cumulative width, and each width parameter r_i controls the "drop" between blocks. Hence expressing shifted Schur functions $S^\star_\mu$ and related α-shifted symmetric functions in terms of the (p_i, r_i) variables will allow us to identity polynomial expansions and a key result relating to the positivity of the shifted Schur function in the falling factorial basis.

3.4 Shifted Schur Functions

In this subsection, we focus on the case $\alpha = 1$; that is, the shifted Schur functions $S^\star_\mu$ discussed here are 1-shifted symmetric functions.

Theorem 3.2 (Alexandersson & Féray, 2017). For every $\mu \in \mathrm{Par}$, the polynomial $S^\star_\mu(\mathbf{r}^p)$ has nonnegative rational coefficients in the falling factorial basis $\mathfrak{B}$.

Proof. Let us fix $\mu \in \mathrm{Par}(n)$. Further, let *S*, *T* be set partitions of [k] with *s* and *t* parts respectively. We shall first express $S^\star_\mu$ as a sum of smaller pieces. We will then prove that each piece is nonnegative in $\mathfrak{B}$. Define the polynomials

$$A^\mu_{S,T}(x_1, \ldots, x_s, y_1, \ldots, y_t) := \sum_{\substack{\sigma,\tau \in \mathfrak{S}_k \\ C(\sigma) \leq S \\ C(\tau) \leq T}} \chi^\mu(\sigma\tau)\varepsilon(\tau) \prod_{\substack{S_i \in S \\ T_j \in T}} x_i^{|C(\sigma)/S_i|} y_j^{|C(\tau)/T_j|}.$$

And define, for $\sigma, \tau \in \widetilde{\mathfrak{S}}_k$, the polynomial $N_{\sigma,\tau}(\mathbf{r}, \mathbf{p})$

$$N_{\sigma,\tau}(\mathbf{r}, \mathbf{p}) := \sum_{\substack{S,T \text{set partitions of}[k] \\ C(\sigma) \leq S, C(\tau) \leq T}} \sum_{\substack{v:S \to [d], w:T \to [d] \\ v,w \text{ compatible and injective}}} \prod_{S_i \in S} p_{v(S_i)}^{|C(\sigma)/S_i|} q_{w(T_j)}^{|C(\tau)/T_j|}.$$

Here $C(\sigma)$ (resp. $C(\tau)$) is the cycle partition of the permutation σ (resp. τ); χ^μ is the irreducible character of $\mathfrak{S}_k$ indexed by μ, and $\varepsilon(\tau)$ is the sign of τ.

We take for granted the fact that we can now express the polynomial

$$S^\star_\mu(\mathbf{r}^p) = \frac{1}{k!} \sum_{\sigma,\tau \in \mathfrak{S}_k} \chi^\mu(\sigma\tau)\varepsilon(\tau) N_{\sigma,\tau}(\mathbf{r}, \mathbf{p}).$$

Using this fact and our definition for $N_{\sigma,\tau}(\mathbf{r}, \mathbf{p})$, we may write

$$S^{\star}_{\mu}(\mathbf{r}^{p}) = \frac{1}{k!} \sum_{S,T\text{set partitions of}[k]} \sum_{\substack{v:S\hookrightarrow[d],w:T\hookrightarrow[d]\\ v,w\text{ compatible}\\ \text{and injective}}} A^{\mu}_{S,T}\left(p_{v(S_1)},\dots,p_{v(S_s)},r_{w(T_1)},\dots,r_{w(T_t)}\right).$$

We now notice that if $A^{\mu}_{S,T}(x_1,\dots,x_s,y_1,\dots,y_t)$ is non-negative in $\mathfrak{B}$ then so is $S^{\star}_{\mu}$, and thus it suffices to show that $A^{\mu}_{S,T}(x_1,\dots,x_s,y_1,\dots,y_t)$ is nonnegative in $\mathfrak{B}$. We appeal to the following Lemmas (see Alexandersson & Féray, 2017, pp. 15–16):

Lemma 3.2 (Alexandersson & Féray, 2017, Lemma 3.5). Fix two set partitions $S=\{S_1,\dots,S_s\}$ and $T=\{T_1,\dots,T_t\}$ of $[k]$, and let $\mu\vdash k$. Then

$$A^{\mu}_{S,T}(x_1,\dots,x_s,y_1,\dots,y_t) = \sum_{\substack{\widetilde{S}\le S\\ \widetilde{T}\le T\\ C(\sigma)\le\widetilde{S}\\ C(\tau)\le\widetilde{T}}} \Big(\sum_{\sigma,\tau\in\mathfrak{S}_k} \chi^{\mu}(\sigma\tau)\cdot\varepsilon(\tau)\Big)\Big(\prod_{S_i\in\widetilde{S}}(x_i)^{|\widetilde{S}/S_i|}\prod_{T_j\in\widetilde{T}}(y_j)^{|\widetilde{T}/T_j|}\Big).$$

Lemma 3.3 (Alexandersson & Féray, 2017, Lemma 3.6) Let $\mu\vdash k$ and let S, T be set partitions of the set $[k]=\{1,\dots,k\}$. Denote by $\mathfrak{S}_S$ (resp. $\mathfrak{S}_T$) the Young subgroup of permutations whose cycles are refinements of the blocks of S (resp. T). Define

$$B^{\mu}_{S,T} = \sum_{\substack{\sigma\in\mathfrak{S}_S\\ \tau\in\mathfrak{S}_T}} \chi^{\mu}(\sigma\tau)\,\varepsilon(\tau)$$

where $\mathfrak{E}$ is the sign character. Then $B^{\mu}_{S,T}\ge 0$.

By the above lemmas, the polynomials $A^{\mu}_{S,T}$ have nonnegative coefficients in $\mathfrak{B}$. Since we previously expressed $S^{\star}_{\mu}(\mathbf{r}^{p})$ as a non-negative linear combination of the $A^{\mu}_{S,T}$, the result follows. □

Shifted Jack Polynomials

4.1 Jack Polynomials

We recall the family of symmetric functions titled the Jack polynomials, denoted by $J^{(\alpha)}_{\mu}$ with $\mu\in\mathrm{Par}$ and parameter α. For $\alpha=1$, Jack polynomials coincide with Schur polynomials up to multiplication by a scalar (Alexandersson & Féray, 2017).

Definition 4.1. We denote shifted Jack polynomials by $J^{\star,(\alpha)}_{\mu}$ (Okounkov & Olshanski, 1997b), where $J^{\star,(1)}_{\mu}$ is a multiple of $S^{\star}_{\mu}$ and the top degree component of $J^{\star,(\alpha)}_{\mu}$ is J_{μ}. We define the shifted Jack polynomials as the unique α-shifted symmetric function of degree at most $|\mu|$ such that for any partition λ with $|\lambda|\le|\mu|$,

$$J^{\star,(\alpha)}_{\mu} = \begin{cases} \frac{H^{(\alpha)}H'^{(\alpha)}}{\alpha^{|\mu|}} & \text{if } \lambda=\mu,\\ 0 & \text{else.}\end{cases}$$

The proof of existence and uniqueness, as well as the top component of $J^{\star,(\alpha)}_{\mu}=J^{(\alpha)}_{\mu}$ is given in Knop and Sahi (1996).

For a diagram λ and $\mu\in\mathrm{Par}$ we define the shifted symmetric function

$$\mathfrak{K}_{\mu}(\lambda) := \begin{cases} n\Big(n-1\Big)\cdots\Big(n-k+1\Big)\frac{K^{\lambda}_{\mu,1^{n-k}}}{K^{\lambda}_{1^n}} & n\ge k,\\ 0 & \text{else.}\end{cases}$$

where K^{β}_{α} are the Kostka numbers. This family has an analogue for parameter α. If we consider $J^{(\alpha)}_{\lambda}=\sum_{\tau\vdash|\lambda|}\widehat{K}^{\lambda,(\alpha)}_{\tau}M_{\tau}$, then we can analogously define for a parameter α the function

$$\mathfrak{K}^{(\alpha)}_{\mu}(\lambda) := \begin{cases} \frac{1}{(n-k)!}\widehat{K}^{\lambda,(\alpha)}_{\mu,1^{n-k}} & n\ge k,\\ 0 & \text{else.}\end{cases}$$

where $\widehat{K}^{\lambda,(\alpha)}_{\mu 1^{n-k}}$ is the coefficient of the monomial $M_{\mu,1^{n-k}}$ in the expansion of $J^{(\alpha)}_{\lambda}$.

Definition 4.2 (Falling Factorial). We define the falling factorial to be the polynomial

$$[x]_n = (x)_n := \underbrace{x(x-1)\cdots(x-n+1)}_{n\text{ factors}} = \prod_{k=1}^{n}(x-k+1) = \prod_{k=0}^{n-1}(x-k).$$

Note we interchangeably use the notation $[x]_n=(x)_n$. Moreover, we note that the falling factorials form a basis $\mathfrak{B}$ for the vector space of polynomials. We provide a brief proof below

Lemma 4.1. Let V_n be the complex vector space of polynomials of degree $\le n$. Then $([x]_0,[x]_1,\dots,[x]_n)$ is a basis of V_n.

This follows by a triangularity argument.

Definition 4.3 (α-Falling Factorial Basis). In the polynomial ring $\mathbb{Q}[\alpha,p_1,\dots,p_d,r_1,\dots,r_d]$ we define

$$\{\,\alpha^{c}\,(p_1)_{a_1}\cdots(p_d)_{a_d}\,(r_1)_{b_1}\cdots(r_d)_{b_d}\,\}_{c,a_1,\dots,a_d,b_1,\dots,b_d\ge 0}$$

to be the α–*falling factorial basis* of this ring. We take for granted that this indeed forms a basis, although the proof is elementary.

Theorem 4.1 (Alexandersson & Féray, 2017, Theorem 1.5). For any $k\in\mathbb{Z}$ the quantity $J^{\star,(\alpha)}_{(k)}(\mathbf{r}^{p})=k!\mathfrak{K}^{(\alpha)}_{(k)}(\mathbf{r}^{p})$ is a polynomial with non-negative rational coefficients in the α-falling factorial basis.

To prove this result, we shall appeal to the following theorem:

Theorem 4.2 (Alexandersson & Féray, 2017). For any integer $k\ge 1$ and any Young diagram λ one has,

$$\frac{1}{k!}J^{\star,(\alpha)}_{(k)}(\lambda) = \mathfrak{K}^{(\alpha)}_{(k)}(\lambda) = \sum_{\substack{A\subseteq\lambda\\ |A|=k\\ \text{column-distinct}}}\Big(\prod_{R\text{ row of }\lambda}P_{|R\cap A|}(\alpha)\Big)$$

where, for $i\ge 0$,

$$P_i(\alpha) := \prod_{j=0}^{i-1}(1+j\alpha)\qquad (P_0(\alpha):=1).$$

Proof of Theorem 4.1, from Alexandersson and Féray (2017). Let us first write

$$\sum_{\substack{A\subseteq\lambda\\ |A|=k\\ \text{column-distinct}}}\Big(\prod_{R\text{ row of }\lambda}P_{|R\cap A|}(\alpha)\Big) = \sum_{A\in\mathscr{CD}(\lambda,k)}w(A)$$

where we define $\mathscr{CD}(\lambda,k)$ to be the set of column-distinct subsets $A\subset\lambda$ of size k and where $w(A):=\prod_{R\text{ row of }\lambda}P_{|R\cap A|}(\alpha)$. Now to each set A we shall associate to it a skeleton $\widehat{A}$ as follows: consider $\lambda=\mathbf{r}^{p}$ in rectangular coordinates, and note that every box in λ lives in a

rectangle $p_i \times r_j$, where we label by i, j respectively the corresponding rows and columns. Then from λwe shall remove any rows and columns whose intersection with A is empty. This shall define our skeleton $\widehat{A}$.

Fix a skeleton $\widehat{A}_0$ and define by a_i and b_i the number of rows and columns labeled i. We observe that A column-distinct $\iff \widehat{A}$ column-distinct, and if two sets $A \neq B$ are such that $\widehat{A} = \widehat{B}$ then $w(A) = w(B)$. From this it follows that:

$$\sum_{\substack{A \in CD(r^p,k) \\ \widehat{A}=\widehat{A}A_0}} w(A) = w(A_0)|\{A \in CD\ (r^p,k)|\widehat{A} = \widehat{A}A_0\}| := \Gamma.$$

It now suffices to check that for each skeleton, Γ has non-negative coefficients in $\mathfrak{B}$. Note that since $w(A_0)$ is a nonnegative polynomial in α we shall focus on the number of sets $A \subset \lambda$ with skeleton $\widehat{A}_0$.

In order to construct such a set A with shape $\lambda = \mathbf{r}^p$, we first decide the distribution of rows and columns. Let us consider $a = (a_1, \ldots, a_i, \ldots),\ b = (b_1, \ldots, b_j, \ldots)$. Then for each i we must pick rows in λ that correspond to the a_i rows in $\widehat{A} = \widehat{A}_0$. We notice that λ has p_i-many boxes labeled i and thus we have $\binom{p_i}{a_i}$ ways to choose a row. By the same logic, there are $\binom{r_i}{b_i}$ ways to choose a column, and thus we have a total of

$$\prod_i \binom{p_i}{a_i} \times \prod_i \binom{r_i}{b_i}$$

many $A \subset \lambda$ of shape $\mathbf{r}^p$ with skeleton $\widehat{A}_0$. We finally observe that $\prod_i \binom{p_i}{a_i} \times \prod_i \binom{r_i}{b_i}$ is non-negative in the basis $\mathfrak{B}$ and hence the result follows. □

A Brief Comparison Between Jack- and Schur-Shifted Functions

While such results establish positivity for shifted Schur functions, corresponding to the $\alpha = 1$ case of Jack polynomials, the situation for general α-Jack polynomials remains more subtle. Determinant-style expressions are known primarily in the case of one-row partitions, $\mu = (k)$, as presented by Knop and Sahi (1996). The extensions of such expressions or positivity results to arbitrary partitions remains not fully understood.

Acknowledgments

We would like to thank Houcine Ben Dali for his challenging and rigorous introduction to the combinatorics of symmetric functions this semester. We are especially grateful for the many helpful conversations that played an increasingly important role in our ability to produce a paper of this kind.

References

Alexandersson, P., & Féray, V. (2017). Shifted symmetric functions and multirectangular coordinates of Young diagrams. *Journal of Algebra*, *483*, 262–305. https://doi.org/10.1016/j.jalgebra.2017.03.036

Aval, J.-C., Féray, V., Novelli, J.-C., & Thibon, J.-Y. (2014). Quasi-symmetric functions as polynomial functions on Young diagrams [Preprint]. *arXiv*. https://arxiv.org/abs/1312.2727

Ben Dali, H. (2025). *Math 290Z: Combinatorics of symmetric functions* [Course notes].

Goulden, I. P., & Jackson, D. M. (1996). Connection coefficients, matchings, maps, and combinatorial conjectures for Jack symmetric functions. *Transactions of the American Mathematical Society*, *348*(3), 873–892. https://doi.org/10.1090/S0002-9947-96-01652-0

Falling factorials form a basis for vector space of polynomials. (2025, April). Mathematics Stack Exchange. https://math.stackexchange.com/questions/961829/falling-factorials-x-0-x-1-ldots-x-n-form-a-basis-for-vector-space-of-po

Lassalle, M. (2009). Jack polynomials and free cumulants. *Advances in Mathematics*, *222*(6), 2227–2269.

Knop, F., & Sahi, S. (1996). Difference equations and symmetric polynomials defined by their zeros. *International Mathematics Research Letters*, *10*, 473–486.

Macdonald, I. G. (1995). *Symmetric functions and Hall polynomials* (2nd ed.). The Clarendon Press, Oxford University Press.

Okounkov, A., & Olshanski, G. (1997). Shifted Jack polynomials, binomial formula, and applications. *Mathematical Research Letters*, *4*(1), 69–78.

Okounkov, A., & Olshanski, G. (1997). Shifted Schur functions. *Algebra i Analiz*, *9*(2), 73–146. (English translation in St. Petersburg Mathematical Journal, 9(2), 239–300, 1998).

Stanley, R. P. (2001). *Enumerative combinatorics (Vol. 2)*. Cambridge University Press.

Stanley, R. P. (2002). Irreducible symmetric group characters of rectangular shape. *Séminaire Lotharingien de Combinatoire*, *50*, eB50d, 11 pp.

Racial Representation and Partisan Dynamics in the U.S. State Legislatures from 1995 to 2000

Todd Zhou
Harvard College '27

This study reexamines the foundational claim that demographic composition, particularly racial makeup, shapes legislative behavior in the United States. Drawing on a novel dataset that I helped digitize from a government census almanac book, which includes 2,630 state legislators across 18 racially salient states from 1995 to 2000, I test whether the percentage of African American residents in a legislative district predicts the roll-call ideology of its representative. The findings challenge conventional wisdom. While bivariate models suggest a strong negative relationship, with Blacker districts electing more liberal lawmakers, this association collapses under minimal statistical conditioning. Once party affiliation, state context, and district socioeconomic characteristics are held constant, racial composition has no substantively or statistically significant effect on roll-call ideology. This null result holds across model specifications, racial thresholds, partisan subgroups, and robustness checks. The one exception, majority-Black Democratic districts, reflects ideological selection rather than policy responsiveness. These results force a reconsideration of both descriptive representation and democratic accountability. In contemporary state legislatures, race determines who wins office but not how they vote. Party discipline, rather than constituency demographics, governs roll-call behavior. Representation has shifted from a delegate model of voter-legislator alignment to a partisan model of ideological conformity. By tracing the structural erasure of demographic influence, this paper offers a revised account of political representation in an emerging era of nationalized partisanship and cautions against equating electoral diversity with substantive policy inclusion.

Introduction

State legislatures are where national ideologies confront the stubborn facts of local demography. Despite having less scholarly prestige than Congress, the nation's 7,383 state legislators collectively wield immense institutional power. They draft and enact far more bills per biennium than their federal counterparts, allocate over $1.5 trillion in public funds, and pilot policy regimes, from environmental deregulation to Medicaid expansion, that routinely cascade upward to the national stage (Carey et al., 2006; Squire, 2012). These areas are not merely provincial echoes of Washington politics, but are foundational to the American policy ecosystem and indispensable to understanding democratic responsiveness at scale.

Yet while political science has made great advances in mapping the ideological terrain of Congress, the micro-foundations of roll-call voting in statehouses are often less known. Roll-call voting, when legislators' individual votes are publicly recorded, offers a rare window into how personal beliefs, party pressures, and constituency interests collide in real time. Yet what drives the policy behavior of state legislators? Two theories offer contrasting answers. One, rooted in democratic representation, casts legislators as delegates of their constituents. In this "constituency-centered" view, electoral accountability binds representatives to local preferences: voters punish deviation, and legislators, in turn, align their voting behavior with district interests (Mansbridge, 1999; Pitkin, 1967). Especially where racial identity is politically salient, as in the post-Voting Rights Act South, the Black population share should, according to this theory, predict more liberal roll-call behavior, consistent with theories of racially linked fate and symbolic representation (Broockman, 2013; Browning et al., 1990; Grofman & Handley, 1991; Welch & Bledsoe, 1985).

A second, increasingly dominant paradigm emphasizes parties, not voters, as the central actors. Party-centered theories describe legislators as agents of ideologically cohesive caucuses—intra-party groups that coordinate strategy, messaging, and collective action within a legislature—whose voting decisions reflect organizational discipline and partisan branding rather than constituency preferences (Levendusky, 2008; McCarty et al., 2006; Poole & Rosenthal, 1997). In this framework, district-level characteristics may affect who wins an election, but exert little influence on how that individual votes once in office. The emergence of strong partisan sorting and elite polarization over the past three decades has only reinforced this interpretation (Carson et al., 2019; Shor & McCarty, 2011).

This study tests two paradigms by analyzing roll-call voting across 18 state legislatures from 1995 to 2000. These legislatures were selected for their demographic relevance, with average Black district shares exceeding 10%. Drawing on the Shor–McCarty Net Promoter-score Framework (Shor & McCarty, 2011), I examine whether the Black population share predicts liberalism in legislators' ideological positions. *The Shor–McCarty NP-score Framework is a cross-state ideological scaling method that places state legislators and members of Congress on a common liberal–conservative dimension using survey and roll-call data*. The analysis is guided by three core research questions:

1. **Unconditional association.** Do districts with higher Black population shares exhibit lower (more liberal) legislator NP scores?
2. **Robustness to context.** Does any such association persist after conditioning on party affiliation and state fixed effects?
3. **Moderation by institutional features.** Are these effects attenuated in more professionalized chambers or among Republican legislators?

Theory and Prior Evidence

Understanding whether and how district demographics shape legislative behavior requires untangling the mechanisms of political

representation and the constraints imposed by institutional structures. Existing research offers competing theories, which I organize here into three thematic strands: the distinction between descriptive and substantive representation, the institutional dominance of parties, and the conditional nature of statehouse dynamics.

1. Descriptive vs. Substantive Representation

At the heart of democratic theory lies a tension between descriptive and substantive representation. Descriptive representation emphasizes shared identity characteristics between voters and elected officials (race, gender, religion, etc.), while substantive representation asks whether those officials advance policies aligned with their constituents' preferences. Mansbridge (1999) identifies four distinct ways that shared identity between representatives and constituents can shape policy responsiveness. First, promissory representation occurs when legislators seek to keep their campaign promises once in office. Second, anticipatory representation involves acting in ways that pre-empt future voter punishment, anticipating how constituents will judge them in the next election. Third, surrogate representation refers to advocating for people outside one's own district who share similar identities or experiences. Finally, gyroscopic representation describes legislators who rely on their own internalized values and sense of duty, rather than external pressure, to guide their decisions, assuming those values align with their constituents.

Empirical evidence supporting these pathways, however, is uneven. Early urban studies, such as Welch and Bledsoe (1985) and Browning, Marshall, and Tabb (1990), found that Black mayors and council members often brought modest but real changes in urban spending priorities, particularly toward social services and minority hiring. Yet as the scope of governance expands from cities to states, these effects become harder to trace. Grofman and Handley (1991), for instance, observe that while the Voting Rights Act increased Black officeholding in southern states, it did not always yield commensurate policy changes. More recent work by Broockman (2013) suggests that Black legislators may exhibit stronger intrinsic motivation to champion racial equity, but their effectiveness is still mediated by institutional rules, partisan caucuses, and the broader legislative environment. If shared racial identity does not robustly predict more liberal voting even in majority-Black districts, then the mechanisms of substantive representation may be fundamentally constrained by other forces.

2. The Partisan Constraint

A powerful counter-narrative to identity-based representation is the growing dominance of parties in structuring legislative behavior. Rooted in spatial theories of roll-call voting, this literature holds that legislators act as agents of their parties, constrained by median voter preferences within their caucus rather than their district (McCarty et al., 2006; Poole & Rosenthal, 1997). In this view, representation is a byproduct of party competition, not constituent alignment. The rise of elite polarization accelerated by gerrymandering, closed primaries, and nationalized media has reinforced this dynamic, rendering party labels increasingly reliable cues for ideological behavior (Levendusky, 2008; Shor & McCarty, 2011).

This logic has two important implications. First, it suggests that the racial or economic composition of a district may influence election outcomes (e.g., which party wins) but not policy outcomes conditional on party control. Once elected, a Democrat from a majority-Black district is expected to vote much like a Democrat from districts with a greater percentage of white residents, regardless of local demographic nuances. Second, it implies that intra-party variation, the range of ideological differences among members of the same political party, is shrinking. The "southern Democrat" or "northeastern Republican" is increasingly a historical artifact, replaced by ideological homogeneity within each party's legislative wing.

Thus, if the observed effect of Black population share on roll-call ideology weakens or disappears after accounting for party and state fixed effects, this suggests that legislators' ideological positions are shaped primarily by partisan alignment rather than by constituency demographics. In other words, once we control for "party," the institutional expression of partisan constraint, the remaining variation reveals how strongly party affiliation channels or overrides demographic influences. In polarized legislatures, this means that racial representation may hold symbolic rather than substantive significance, as party incentives largely determine legislators' policy behavior.

3. Why States Could Differ

Nonetheless, state legislatures are flexible, and it would be unwise to dismiss the potential for constituency influence without considering institutional variation. As Squire (2012) shows, statehouses differ dramatically in their levels of professionalization, defined by salary, staff support, session length, and legislative capacity. In low-professionalized chambers (e.g., in Arkansas and Montana), legislators often hold other jobs, face fewer procedural restrictions, and operate in a looser partisan environment. Such conditions may reduce the effectiveness of party discipline and open the door to more district-responsive behavior.

Electoral rules, too, play a critical role. States with term limits may incentivize short-term constituency service over long-term party loyalty, especially in members' final sessions (Carey et al., 2006). Conversely, states with multi-member districts, nonpartisan primaries, or public financing regimes may induce broader electoral coalitions, potentially heightening responsiveness to racial or economic heterogeneity within districts.

Moreover, the salience of race itself varies by region and historical legacy. In Deep South legislatures with a history of racial bloc voting (where voters of different racial groups consistently support opposing candidates) and constrained Black access to power, descriptive representation may carry different symbolic and mobilizational consequences than in racially diverse coastal states. Studies by Grose (2005) and Batista (2020) suggest that even when Black legislators represent Black-majority districts, their roll-call behavior is shaped more by institutional context than by co-ethnic solidarity alone. These conditional dynamics underscore the need for a multi-level empirical approach: one that tests not only for aggregate relationships between racial composition and ideology, but also for heterogeneity across states, chambers, and partisan environments.

Methods

This paper evaluates whether racial constituency composition (specifically, the percent of African American residents in a state legislative district) predicts the roll-call ideology of the elected representative. To do so, I constructed a novel, legislator-level dataset using the Gov 2474's Optical Character Recognition (GOCR) algorithm. I developed this AI algorithm extract tables and numerical data from scanned or photographed pages of physical books. It merges

roll-call ideal points with detailed socioeconomic and demographic attributes of each legislator's district. The resulting data portrait offers an unusually fine-grained lens on constituency-legislator linkages across 18 states during a formative period of American polarization (1995–2000). The empirical strategy proceeds in three phases: (1) documenting raw descriptive relationships, (2) estimating causal effects via fixed-effects regression models, and (3) probing the stability of findings across institutional and partisan contexts.

Unit of Observation and Dataset Structure

The dataset consists of 2,630 unique legislators observed between 1995 and 2000, each represented by a single row reflecting five-year averages of both outcome and covariates. This collapsed cross-section minimizes volatility due to short-term shocks or legislative turnover while preserving meaningful variation across districts and states. By focusing on incumbents only, I ensure that the ideological measures reflect actual voting behavior rather than campaign signaling or appointment bias. I then combine this dataset with the Shor and McCarty (2019) dataset to receive all the NP scores (discussed in the next section) calculated for each representative in each district.

The sample includes both chambers (85% House, 15% Senate) and both major parties. States were selected if their average district-level Black population exceeded 10%, a criterion that ensures sufficient racial variation while anchoring the analysis in geographies where race plausibly shapes electoral incentives. Five states—Georgia, Texas, North Carolina, Louisiana, and Mississippi—account for 26% of the sample and contain nearly all majority-Black districts, making geographic balance and fixed-effects particularly vital. This was made possible with the advent of OCR algorithms to help digitize this robust data for in-depth analysis.

Measuring Ideology: The NP Score

The dependent variable is the NP Score, a continuous ideal-point estimate developed by Shor and McCarty (2011) using state-level roll-call data. It places each legislator on a latent ideological spectrum anchored by party caucus medians, scaled to allow comparison across states and chambers. NP scores in the dataset range from –2.65 (very liberal) to 2.35 (very conservative), with a mean near zero and a standard deviation of 0.78. The inter-party difference is vast: Democrats average 0.39 while Republicans cluster at 0.83, a partisan gap that makes within-party variance seem negligible (SD = 0.42 for Democrats, SD = 0.34 for Republicans).

This strong polarization introduces both an empirical challenge and an inferential opportunity. Because party affiliation predicts ideological position, meaningful estimation of racial constituency effects must occur within-party. Accordingly, all models are conditioned on party or are estimated in split samples to avoid mistaking party sorting for constituency responsiveness.

Focal Independent Variable: Percent African American

The primary explanatory variable is the percentage of African American residents in a legislator's district, derived from the 2000 U.S. Census and joined to legislative boundaries. This variable is sharply skewed: roughly two-thirds of districts fall below 15% Black, while a concentrated right tail includes a substantial number of districts above 50%, and a handful above 80%. This bimodal distribution is quite fortunate. It enables comparisons across starkly different racial contexts while preserving a broad mid-range for detecting marginal effects.

A naive bivariate regression suggests a dramatic relationship: NP scores fall nearly 1.5 units when moving from the bottom to the top decile of district Black share. Yet this pattern almost certainly reflects compositional factors (party, geography, and socioeconomic context) rather than a direct causal effect. Since the distribution is analytically central, we are able to discern nonlinear patterns (e.g., thresholds or saturation points), conduct stratified modeling (the 10–50% "competitive" band), and perform robust tests of responsiveness across the full range of demographic variation.

Covariates and Fixed Effects

Multivariable models included: average income (per $1,000) to capture socioeconomic status, college attainment, to control for educational composition, party affiliation, and state fixed effects, to absorb unobserved institutional, regional, or partisan differences.

Together, these controls isolated the independent effect of race from correlated social, economic, and geographic factors. State fixed effects in particular were essential, as several high-Black-share districts clustered in Southern states with sharply distinct political cultures. Robust standard errors were clustered by state throughout, accounting for intra-state correlation and variance heterogeneity.

Modeling Strategy

We conducted the following regression analyses to better understand the phenomenon.

A. Bivariate Ordinary Least Squares (OLS)

Initial models regressed the NP score on the percent Black without controls. These regressions established raw associations and match prevailing assumptions in the literature. Both pooled and party-stratified models were estimated.

B. Multivariable OLS with Fixed Effects

Subsequent models added income, party, and state dummy variables. These tested whether the racial coefficient survives once confounding variance was absorbed. The fully specified model was:

$$NP_i = \beta_0 + \beta_1 BlackShare_i + \beta_2 Income_i + \beta_3 Party_i + \gamma_s + \varepsilon_i$$

where:

- NP_i is legislator i's ideological (NP) score,
- $BlackShare_i$ is the percentage of i's district that is African American,
- γ_s is the state fixed effects, and
- ε_i is the heteroskedasticity-robust error term, clustered by state

This specification tested the core hypothesis of whether district racial composition independently shapes legislative ideology after controlling for party and geography.

C. Band-Restricted Models (10–50%)

Because the racial extremes (less than 10%, greater than 50%) dominate the bivariate slope, models were also estimated within the 10–50% Black range, where racial composition may generate more cross-group electoral competition. This subset comprised moderate, interracial districts where policy responsiveness should be highest.

D. Per-State Regressions

To assess local robustness, separate OLS regressions were run within each state. This allowed the recovery of state-specific slopes and a test of whether the null finding reflects an average over diverging subnational trends. In most states, bivariate slopes remained negative, but these models did not control for the party, and their attenuation in multivariable settings further reinforced the central findings.

E. Diagnostic Tests

Model integrity is assessed via:

- Variance inflation factors (VIF < 1.5) to rule out multicollinearity,
- Oster coefficient stability tests ($\delta^* = 1.46$) to assess omitted variable bias,
- Moran's I on residuals to test for spatial autocorrelation ($I = 0.015$, $p = .34$).

We utilized the inspection of LOWESS curves (locally weighted regressions that fit a smooth line through scatterplot data, which help reveal nonlinear patterns), scatterplots, and residual distributions for functional form and heteroskedasticity analysis.

Overall, this design allowed for a clean test of the competing models of legislative behavior. If a constituency-centered model holds, we should observe a robust, negative relationship between district Black share and liberalism, even within party and within state. If a party-centered model holds, the racial slope should collapse under control, and ideology should be driven by partisan identity rather than demographic context.

By estimating each model with layered controls, restricted bands, and disaggregated subpopulations, the empirical strategy revealed where and when race matters, and where its influence was entirely dulled or nullified.

Results

1. Descriptive Patterns: Strong Raw Associations

The bivariate association between district racial composition and legislator ideology is striking in magnitude. Across 2,630 legislators from 18 Black-salient states between 1995 and 2000, a simple OLS regression of NP score on percent Black yields a highly significant negative slope: every additional percentage point in Black population correlates with a 0.010 decrease in NP score ($p < .001$, $R^2 = 0.168$). The ideological difference between a district at the 10th percentile of Black share (≈ 2%) and one at the 90th percentile (≈ 52%) exceeds 1 full NP unit, equivalent to moving from a centrist Democrat to a committed progressive.

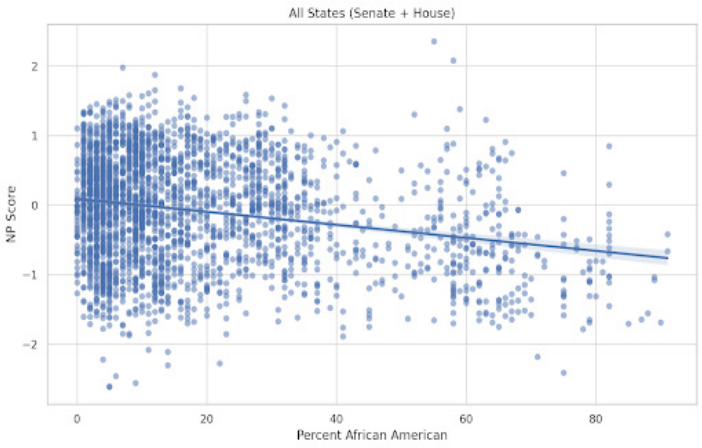


Figure 1. NP Scores vs. Percent African American in Each District

Party-stratified bivariate models, however, appear to support competing interpretations. Among Democrats, the slope remains negative ($\beta = -0.0039$, $p < .001$), suggesting that Blacker districts elect more liberal Democrats. Among Republicans, the sign flips ($\beta = +0.0051$, $p < .001$), indicating a mild conservative shift in more racially diverse constituencies.

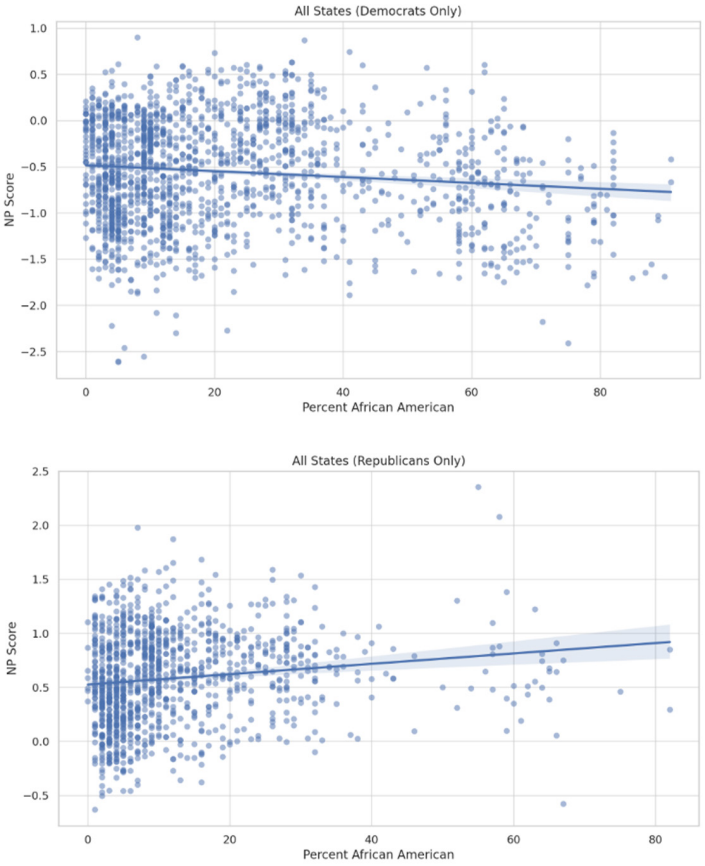


Figure 2. Percent African American and Relative NP Score Separated by Political Party

Yet the low R^2 values in both cases (~.02) signal that race explains little of the within-party variance, prompting a deeper investigation.

2. Conditioning on State and Party: The Collapse of Constituency Effects

The strong negative relationship between district Black population share and legislator ideology seen in the pooled data collapses under even minimal statistical conditioning. Introducing state fixed effects, a necessary adjustment given the geographic clustering of majority-Black districts in the Deep South and urban Rust Belt, helps attenuate the apparent relationship and shifts its substantive interpretation.

The solid orange line represents the simple bivariate OLS relationship, showing a slope of about $\beta = -0.009$. After adding state fixed effects (dashed red line), the slope becomes slightly steeper, at $\beta = -0.012$. This means that, within individual states, districts with larger Black populations tend to elect somewhat

more liberal legislators, even more so than the pooled estimate suggests. Visually, the fixed-effects line looks flatter in the middle of the distribution and bends downward only in the far right tail (above roughly 50% Black population). This aspect of the model reflects where observations are most concentrated, not a weaker overall effect. Statistically, including state fixed effects actually increases the magnitude of the relationship. In other words, once we account for differences between states (such as the fact that liberal, heavily Black districts often coexist with conservative, mostly white states), the link between race and ideology not only persists but becomes sharper within states. This indicates that the connection is rooted in in-state demographic variation, not merely in cross-state political clustering.

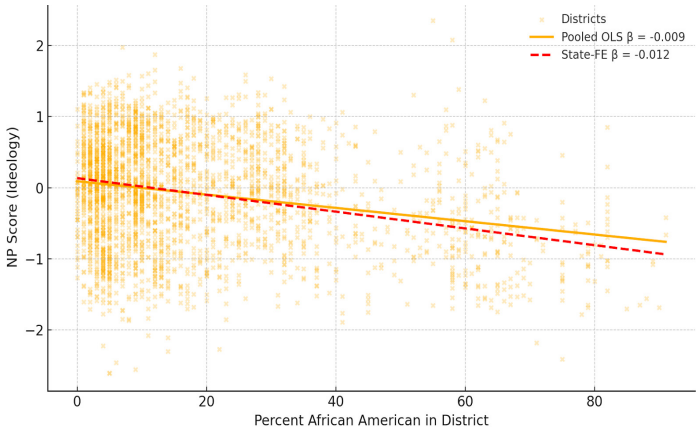


Figure 3. Percent African American Districts and NP Score with Pooled OLS and State Fixed Effects

When party affiliation is introduced into the model alongside state fixed effects and socioeconomic controls, the racial coefficient shrinks further, and in most cases disappears entirely. In the districts that have a 10-50% Black population (which includes the majority of electorally competitive, racially mixed districts), the multivariable regression finds that a 10-point increase in Black population share raises a Democrat's NP score by 0.023 (p = .260), and raises a Republican's NP score by +0.014 (p = .518).

However, these slopes are statistically indistinguishable from zero, and their visual counterparts appear almost perfectly flat. Thus, the explanatory power of race is minimal: within this critical middle band, constituency demographics exert no meaningful influence on how legislators vote.

Switching from Democrat to Republican increases a legislator's NP score by 1.17 units, more than 25 times the ideological impact of a 20-point increase in Black population for a Democrat. The disparity in magnitude is so extreme that it renders the racial slope not only statistically negligible but practically irrelevant in predictive terms.

Together, these findings decisively reject the notion that racial constituency composition independently drives roll-call ideology in modern state legislatures. Whatever ideological variation is attributed to racial demography in raw models is fully absorbed by partisan sorting and geographic structure. Once those filters are applied, constituency race explains almost nothing about legislative behavior.

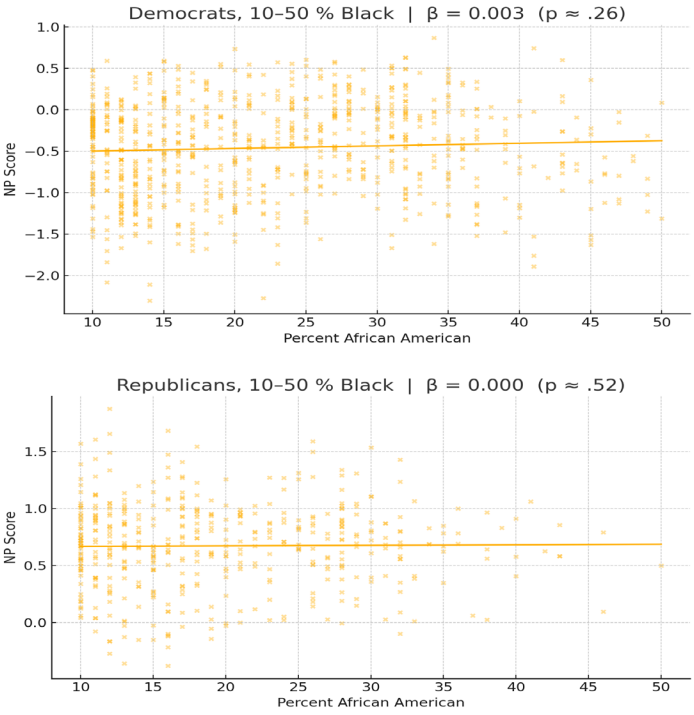


Figure 4. 10-50% Black Districts and Relative NP Score Separated by Political Party

3. The Role of District Composition Extremes

The only setting where race meaningfully predicts ideology is at the tails of the racial distribution, particularly in majority-Black Democratic districts. LOWESS plots show a pronounced downward inflection in Democratic NP scores beginning around 25% Black share. Among the 248 majority-Black Democratic districts in the dataset, the fitted value approaches -1.2 NP units, which is substantially to the left of the caucus mean. This supports the descriptive representation thesis: symbolic and substantive alignment may co-occur in Black-controlled constituencies, but only in safely Democratic environments where primary voters dominate.

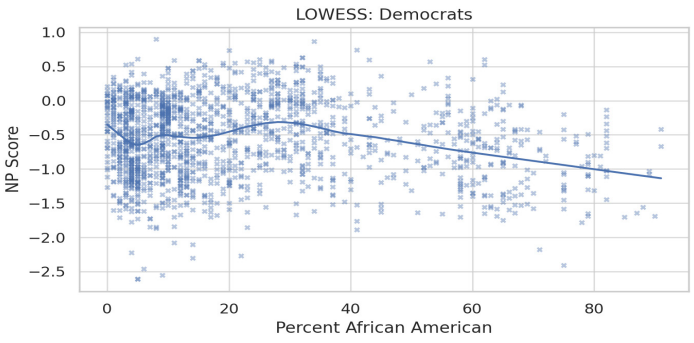


Figure 5. LOWESS Plot of Percent African American Districts and NP Score for Democrats

For Republicans, however, the pattern is reversed but muted. Between 0–25% Black share, there is a mild increase and plateau in conservatism, consistent with reactive polarization or racial threat theories in the literature review. Yet the tail is thin: only seven Republican legislators represent districts with greater than or equal to 50% Black residents, limiting generalizability. Moreover,

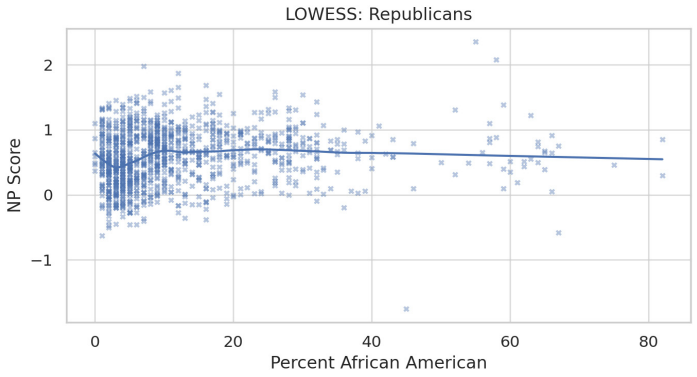


Figure 6. LOWESS Plot of Percent African American Districts and NP Score for Republicans

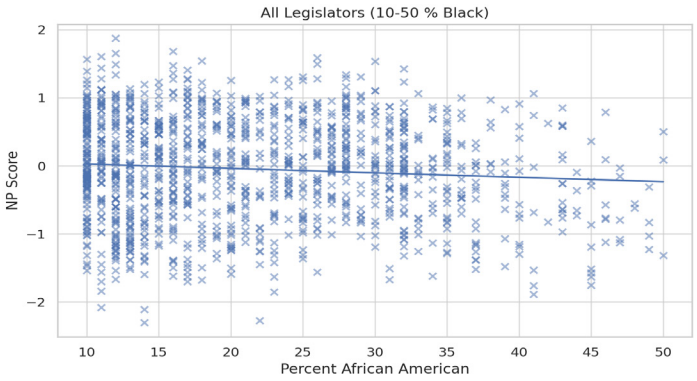


Figure 7. 10% to 50% Black Districts and relative NP scores (Across Parties)

once the Black population share exceeds about 30%, almost no Republican legislators remain in the sample. This highlights how strong partisan sorting by race limits meaningful comparison between parties, since districts with large Black populations are overwhelmingly represented by Democrats.

This pattern implies that racial composition influences roll-call ideology only under conditions of partisan insulation, and even then, its effect is confined to a narrow and structurally exceptional subset of districts.

Among Democrats, meaningful racial responsiveness emerges only once Black residents constitute an electoral majority. In these majority-Black districts, where Democratic incumbents face minimal general election threat and are accountable primarily to Black primary electorates, legislators tend to adopt more liberal roll-call positions. This suggests that descriptive and substantive representation can align, but only in settings where electoral dynamics are stable, homogeneous, and intra-party. In effect, race matters when it dominates the political environment and the Democratic party allows space for ideological expression consistent with Black constituent preferences.

Among Republicans, the story is different. The slight increase in conservatism as Black share rises in predominantly white districts likely reflects defensive ideological positioning, where GOP legislators respond to perceived racial threat. However, the absence of a surplus of Republicans in majority-Black districts means we cannot observe whether similar responsiveness would emerge under reversed partisan conditions. The data simply do not contain many of those special cases, due to strong partisan sorting that prevents Republicans from holding diverse or Black-majority seats.

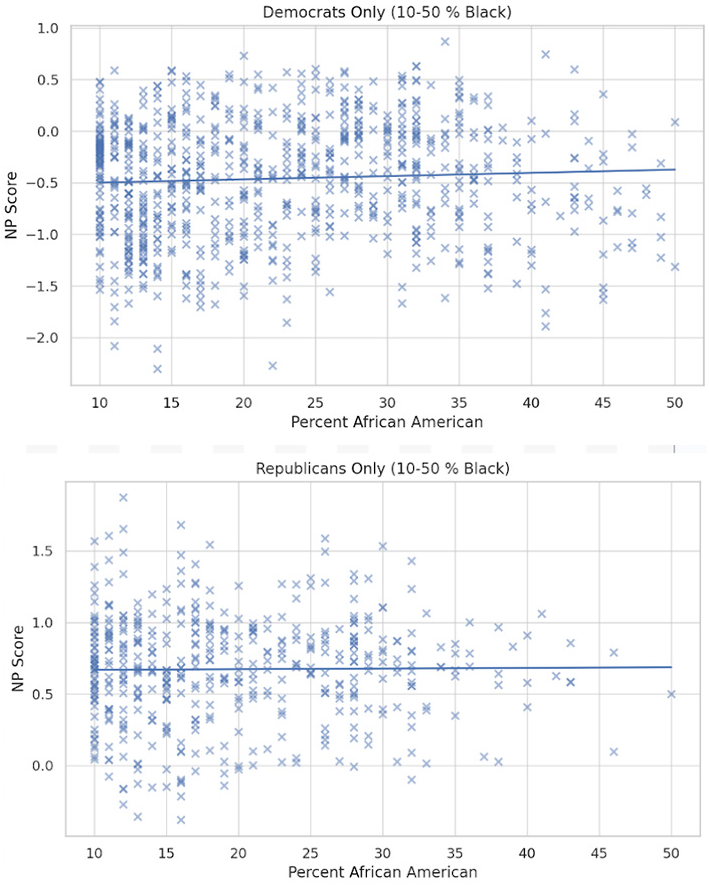


Figure 8. 10% to 50% Black Districts and relative NP scores separated by Political Party

4. Intermediate Districts: Ideological Convergence

Districts within the 10–50% Black population range, often viewed as contexts where interracial coalition politics could emerge, show virtually no ideological responsiveness. In this competitive band, the overall slope is –0.0065 (p = .029), vanishing once income and party are added.

Next, we separated the effect for Democrats and Republicans. Surprisingly, among Democrats, the coefficient was **positive** (β = 0.0031, p = .180).

One plausible interpretation is that moderate Democrats disproportionately represent racially mixed or modestly Black districts, especially in the rural South and parts of the Midwest. These Democrats are often electorally viable precisely because of their ideological moderation and cross-racial appeal. As such, even in districts with sizable Black populations (e.g., 20–40%), Democratic incumbents may retain centrist voting records in order to balance racially heterogeneous coalitions or appease conservative white voters. The result is a flattening or mild inversion of the race–ideology relationship within the party.

This points to a broader structural phenomenon: ideological diversity within the Democratic Party is not randomly distributed but is instead geographically and demographically patterned. In the South, for instance, Black voters are essential to Democratic coalitions but often reside in districts that also include white rural conservatives. In such districts, the electoral incentive may not

be to respond more strongly to Black constituents, but to be moderate enough to hold the seat. Hence, the positive slope may reflect electoral strategy, not unresponsiveness.

In contrast, among Republican legislators, the coefficient is flat ($\beta = 0.0005$, $p = .86$), with NP scores tightly clustered between 0.55 and 0.75 regardless of district racial composition. This ideological compression reflects the strong internal cohesion and nationalization of the GOP caucus, where variation in local demographics appears to have virtually no influence on roll-call behavior. Even in more racially diverse districts, of which there are few, Republicans behave ideologically like their counterparts in overwhelmingly white areas.

This pattern is consistent with theories of partisan insulation and reactive polarization. In an era of strong party discipline, Republicans who survive in diverse districts are typically those who already conform to caucus ideological norms. Rather than adapting to constituent preferences, they represent a filtered selection: either because moderate Republicans cannot win nomination primaries, or because pressure from both sides forces them to double down on ideological identity as a form of signaling their partisanship.

5. Subnational Context: Where Race Still Matters

When the analysis is disaggregated to the state level, the core pattern remains broadly consistent: district racial composition is a weak and inconsistent predictor of roll-call ideology once partisan control is taken into account. That said, modest residual effects persist in a subset of states, particularly those with a history of racially polarized voting and minimal partisan competition.

In the Deep South, states like Mississippi ($\beta = -0.016$), South Carolina ($\beta = -0.019$), and Alabama ($\beta = -0.013$) continue to exhibit negative slopes even in within-state bivariate regressions. These contexts are characterized by high levels of racial segregation, institutionalized bloc voting, and durable majoritarian districting practices that often carve out Black supermajorities in otherwise Republican-dominated environments. In such settings, Black constituents may still exert influence over ideological outputs, but only in safe Democratic seats where primary elections are decisive and general election turnover is rare.

Similarly, Rust Belt states like Illinois and Ohio register sizable negative race–ideology coefficients. These effects are driven by sharp intra-state splits between urban minority-majority districts (often electing progressive Democrats) and rural or ex-urban white districts, invariably sending conservative Republicans. However, this contrast reflects not a continuous representational mechanism, but geographic sorting and partisan polarization. Once again, the effect of race is entirely confounded with party control, and within-party slopes attenuate or vanish under control.

Even in states where race retains statistical significance, its magnitude and explanatory power remain modest. In no case does racial composition explain more than 30% of variance in ideology and most states cluster well below that. The strongest effects occur precisely where racial and partisan boundaries coincide, suggesting that the observed relationships reflect electoral geography and seat allocation, not responsiveness to constituency demographics per se.

By contrast, in fast-growing and demographically dynamic states like Florida, Georgia, and North Carolina, the racial slope is statistically indistinguishable from zero. These states are experiencing rapid suburban diversification and increasingly nationalized campaigns, yet ideological behavior remains fixed along party lines. In such environments, racial demography has become too fluid and partisan incentives too rigid to sustain constituency-driven ideological adaptation.

This contrast is analytically revealing. Where race still "matters" for roll-call behavior, it does so only in structurally exceptional contexts: where Black voters dominate primary electorates, party switching is rare, and general elections are uncompetitive.

NOTE: Model diagnostics support the robustness of these results. Variance inflation factors (VIFs) remain below 1.5 for all covariates, dispelling concerns of multicollinearity. Oster bounds further indicate that unobserved confounding would need to be implausibly large ($\delta^* = 1.46$) to overturn the main null findings. Spatial autocorrelation is negligible (Moran's $I = 0.015$, $p = 0.34$), suggesting no residual spatial clustering of errors.

6. Synthesis of Findings

Taken together, the model results and visual diagnostics present a more direct picture: district racial composition has limited explanatory power for legislative ideology, except in the most racially polarized and electorally lopsided contexts. In politically competitive or demographically mixed districts (where constituency influence should, in theory, be strongest) legislative behavior remains largely unresponsive to racial composition. Party affiliation, not district demographics, overwhelmingly structures ideological outcomes.

The findings point toward a model of bounded responsiveness. In majority-Black, safely Democratic districts, racial congruence aligns with more liberal roll-call behavior, consistent with theories of descriptive representation. In predominantly White Republican districts, racial diversity has a modest amplifying effect on conservatism. Yet in the ideological median, across districts where partisan control is contested or racial composition is moderate, the signal from racial demographics is effectively muted and absorbed entirely by partisan sorting.

Robustness Checks and Extensions

The central finding of this study, that district racial composition has no independent effect on state legislator ideology once party and geography are held constant, is striking. Given the theoretical weight this claim carries, it demands rigorous scrutiny. Accordingly, this section puts the core models to an array of robustness checks and analytical extensions designed to address concerns of specification sensitivity, measurement validity, and contextual scope.

1. Per-State Regressions: Apparent Effects without Party Control

One initial concern is that the pooled models might mask heterogeneity across states. Perhaps race matters in some institutional contexts but not others. To test this, we estimate separate bivariate regressions for each of the 18 states. These state-specific models consistently yield negative slopes (i.e., higher Black share is associated with more liberal roll-call behavior).

At face value, this might seem to contradict the core null hypothesis. However, these state-level regressions do not include party controls, and the patterns they reveal are entirely consistent with compositional sorting. That is, heavily Black districts elect Democrats, and Democrats vote more liberally, producing a

race–ideology link that disappears once intra-party comparisons are introduced. When multivariable models are run within states and within parties, the race coefficient attenuates to zero or reverses. Thus, these per-state slopes reinforce a core argument: race predicts electoral outcomes, not legislative behavior.

2. Functional Form and Nonlinearity

Given the skewed distribution of percent Black and the non-monotonic shape of LOWESS plots, I test several alternative functional forms:

- Decile dummies for racial share reveal that ideological differences are concentrated in the highest two deciles (>50% Black). Below that, the slopes are flat or positive.
- Piecewise linear splines with knots at 25% and 50% confirm that the relationship steepens only in majority-Black districts and only among Democrats.
- Log-transformed racial share yields no improvement in fit or significance.

There is no evidence of a smooth, continuous relationship between racial demography and ideology. Whatever responsiveness exists is confined to extreme racial environments, where party control is already locked.

3. Intermediate Band: Where Accountability Should Be Strongest

The core theoretical test of constituency influence lies in racially mixed, electorally competitive districts, those with 10–50% Black population. This is where interracial coalitions are plausible, primary electorates are diverse, and legislators face electoral cross-pressures.

In this band, we observe that the bivariate slope is shallow ($\beta = -0.0065$) and vanishes under controls. Moreover, among Democrats, the slope is slightly positive (0.0031), suggesting that moderate Democrats often win in racially integrated seats. Among Republicans, the slope is near zero (+0.0005), with NP scores tightly clustered.

These nulls are not incidental. They appear precisely where theories of electoral sanctioning and voter-legislator congruence would predict responsiveness to be strongest. Their absence suggests that the expected link between who voters are and how legislators behave fails to materialize, precisely in the contexts where electoral accountability should be most active.

4. Outlier and Leverage Tests

To assess whether the core findings are driven by a small number of extreme cases, a series of outlier and influence diagnostics were conducted. First, I re-estimate the full specification after excluding Mississippi and South Carolina, states that contribute a disproportionate share of high-Black-population districts. The results remain substantively and statistically unchanged. Second, I remove all districts with either greater than 80% or less than 2% Black population to test for nonlinear leverage at the demographic extremes. Again, the estimated coefficients remain stable. Finally, I implement robust regression models using Huber weights to downweight high-leverage observations. These models confirm that no individual district or small subset disproportionately influences the results. Taken together, these tests indicate that the core null findings are not artifacts of outliers, leverage, or heteroskedasticity, but rather reflect a consistent pattern across a wide range of institutional and demographic contexts.

5. Spatial and Institutional Diagnostics

To test for spatial dependence in the residuals, I compute Moran's I on the fixed-effects residuals. The result ($I = 0.015$, $p = 0.34$) indicates no significant spatial autocorrelation, suggesting that unobserved geographic clustering does not bias the results. Additionally, I explore whether institutional variation moderates the relationship between racial composition and ideology. Interaction models incorporating chamber type (House vs. Senate), legislative professionalization, and the presence of term limits reveal no systematic heterogeneity. Even in structurally weaker legislatures, settings often assumed to create greater responsiveness to local constituencies, the racial composition coefficient remains null. These results further reinforce the broader claim: party affiliation and caucus discipline now exert overwhelming influence on legislative ideology, even in institutional contexts that historically allowed greater latitude for constituency-driven behavior.

Across every diagnostic frontier, the conclusion remains consistent: district racial composition does not meaningfully shape legislator ideology once party and geography are accounted for. Whatever role racial demography may have once played in legislative behavior has been absorbed by partisan structures, constrained by institutional design, and filtered through nationalized coalitional politics. This robustness confirms that the observed null is not a modeling accident, but a structural reality: representation in the modern statehouse is partisan by nature and demographically stagnant by design.

Discussion and Implications

The empirical results presented above challenge a foundational intuition in democratic theory: the social composition of a constituency, especially its racial makeup, should meaningfully constrain the behavior of its representative. Nowhere has this idea been more normatively and empirically central than in the literature on descriptive representation, which holds that co-ethnic elected officials are more likely to translate group preferences into policy outputs (Broockman, 2013). Yet the evidence from 2,630 state legislators across 18 racially salient states suggests that this representational linkage is fragile, conditional, and structurally downsized by partisan control.

1. Descriptive Representation as Electoral Gatekeeping

One of the most striking patterns in the data is the conditionality of racial effects. In majority-Black districts, Democratic legislators are indeed more liberal: NP scores fall sharply once Black share surpasses 50%, especially in southern states. This is textbook descriptive representation: a racially cohesive electorate selects candidates who are both symbolically and substantively aligned with group interests (Broockman, 2013; Mansbridge, 1999). But that effect is limited to precisely the subset of districts where electoral controversy is minimal. That is, these are seats where the general election is a formality and where Democratic nominees emerge from racially homogenous primary electorates.

In all other settings, including the competitive middle of the racial distribution (10–50% Black), descriptive representation fails to predict ideological behavior. The implication is that race matters as a gatekeeping variable (determining who gains office) but not as an ongoing influence on what legislators do once there. This flips the causal direction often assumed in representation theory. Rather than constituency demographics shaping legislative action,

they select candidates already aligned with prevailing caucus norms. The representational "effect" of race is thus absorbed upstream (in recruitment, nomination, and selection) and not downstream in voting behavior.

2. Party as the Dominant Interpretive Lens

If race predicts who wins, the party explains what they do. The effect of partisan identity on NP scores makes all other variables insignificant in magnitude and impact. This finding aligns with a growing literature on party-centric representation, which contends that roll-call behavior in polarized systems is structured not by individual preferences or local contexts, but by the ideological demands of legislative caucuses (Levendusky, 2008; McCarty et al., 2006; Shor and McCarty, 2011).

In this view, parties are not just aggregators of preferences; they are gatekeepers of ideological behavior, imposing discipline, controlling committee assignments, and coordinating electoral support. Once elected, legislators are institutionally included in partisan hierarchies that reward loyalty and penalize deviation. The result is a form of "programmatic homogeneity" that renders district-level variation constant.

3. The Geography of Responsiveness: Narrow and Receding

The one domain where constituency effects still operate (majority-Black Democratic districts) has clear limits. These districts are geographically concentrated, primarily in the Deep South and a handful of northern urban cores, and institutionally protected by Voting Rights Act-era districting rules, which are the exceptions that demonstrate the rule.

Outside of these districts, even where Black shares rise substantially, ideological behavior remains flat, especially among Republicans who exhibit ideological convergence across a wide range of demographic settings. The mild upward slope in NP scores in GOP districts with rising Black share likely reflects selective survival: only the most conservative Republicans can win in such districts, reinforcing caucus norms rather than diluting them.

This spatial boundedness of representational responsiveness has real consequences. It suggests that race continues to shape representational outcomes only where partisan risk is low. Where competition exists (where accountability should matter most), the ideological constraints imposed by parties crowd out demographic responsiveness. Thus, it is evident that it may be the maps and the parties who draw them that define ideological space.

4. Accountability After Polarization: A Broken Feedback Loop

Perhaps the most profound implication of these findings lies in their challenge to the idea of democratic accountability itself. Classic models propose a feedback loop: voters express preferences, then representatives respond, and then voters reward or punish accordingly. But the evidence here suggests that this loop has been broken in the sector of racial representation.

Legislators do not appear to adjust their roll-call behavior in response to changes in racial composition unless they are running in heavily homogenous, safe seats. In mixed districts, where the risk of electoral punishment is real, legislators conform not to district preferences but to party position. This finding not only echoes the congressional literature on asymmetric polarization and partisan sorting but also adds a new layer: even at the subnational level, where heterogeneity is greater and professionalization is lower, representation has become structurally unresponsive to demographic diversity.

This reframing warrants greater consideration of the chain that links voters to policy. If constituents cannot shape legislative behavior through demographic presence except under rare, homogenous conditions, then what remains of the case for descriptive representation? One possible answer is that demographics shape partisanship, and partisanship, in turn, shapes representation: voters choose parties that seem to reflect their group's interests, and those parties carry those priorities into government. Yet the evidence here suggests that this connection has weakened, or even reversed. Demographics still influence who gets elected, but once a party holds power, those same identities no longer majorly influence how its legislators behave. In effect, parties have absorbed and standardized the representational work that demographics once performed. When party loyalty overpowers local diversity, and demographic differences no longer produce meaningful policy variation, the link between citizens and their representatives becomes moderated through institutions whose incentives serve parties first and voters only second.

Limitations

As with any observational analysis, several caveats apply. First, the dataset is restricted to incumbents and to states with racially salient electorates. While this allows for targeted insight, it may limit external validity to less diverse or more institutionally distinct contexts. Second, NP scores, while widely used and validated, are a one-dimensional proxy for ideology. They may obscure multidimensional representational behaviors (such as constituent services, agenda-setting, or rhetorical framing) that fall outside the roll-call record. Third, this study is confined to the late 1990s and early 2000s. This temporal window is analytically useful, but it may not fully reflect subsequent changes, including the intensification of nationalized politics, the collapse of rural Democratic strongholds, and the rise of "identity caucuses" within party coalitions.

Conclusion

This study set out to examine whether the racial composition of state legislative districts in the United States shapes the ideological behavior of their elected representatives. Drawing on an original, five-year dataset of over 2,600 state legislators across 18 racially salient states, the answer that emerges is both empirically robust and theoretically unsettling: district race matters profoundly for electoral outcomes, but minimally, if at all, for legislative behavior once partisan control is established.

At the descriptive level, racial demography appears to predict ideology: Blacker districts elect more liberal legislators. Yet this bivariate association dissolves under even modest controls. State fixed effects, district income, and especially party affiliation absorb the signal. Within-party analyses reveal that, in the overwhelming majority of cases, the percentage of Black residents has no statistically or substantively significant effect on how legislators vote. The exceptions (majority-Black Democratic districts in the Deep South and inner cities) prove this trend: race influences policy only where electoral homogeneity and partisan insulation already preclude competition.

What emerges, then, is a framework of partisan dominance. Legislators in this period may not have been ideological delegates of their districts' racial or economic composition, but rather agents of

their caucus, shaped by party medians and institutional incentives. Representation, even in the relatively decentralized environment of U.S. statehouses, appeared to shift upward, away from local and citizen preferences and toward centralized party control. While these findings are drawn from the late 1990s and early 2000s, the dynamics they reveal continue to raise important questions about how partisan structures mediate the relationship between constituents and their representatives today.

In a period marked by renewed debate over race, representation, and democracy, this project offers a cautionary but clarifying insight: descriptive presence is not and may never have been a sufficient condition for policy influence. Constituency demography remains electorally powerful, but its representational leverage is structurally contingent, activated only when partisan incentives allow it. The next step is not merely to ask who holds office, but to understand what makes representation substantive rather than symbolic. Future research should investigate how party networks, agenda control, and access to institutional resources determine whether identity in office translates into policy action. Comparative and historical analyses can clarify when and under what political conditions identity-based coalitions transform representation into durable change.

The aim, then, is not simply to mirror demographic diversity, but to design institutions that make that diversity meaningful. Descriptive representation alone cannot guarantee policy influence. It becomes substantive only when the structures surrounding it enable translation into power. Moving forward, the task is to reconnect legislators to the social realities of their constituencies, where identity informs but does not predetermine political behavior. By examining how parties, agendas, and electoral rules mediate these connections, future work can illuminate pathways toward a more accountable and adaptive model of democracy, one where inclusion is not merely a matter of presence, but a genuine capacity to shape outcomes.

Acknowledgements

The author would like to thank Professor James M. Snyder, the Leroy B. Williams Professor of History and Political Science at Harvard University, for his wonderful guidance and support during the entire project.

References

Broockman, D. E. (2013). Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives. *American Journal of Political Science*, *57*(3), 521–536.

Browning, R. P., Marshall, D. R., & Tabb, D. H. (1990). *Racial Politics in American Cities*. New York: Longman.

Grofman, B., & Handley, L. (1991). The Impact of the Voting Rights Act on Black Representation in Southern State Legislatures. *Social Science Quarterly*, *72*(4), 867–893.

Grose, C. R. (2005). Disentangling Constituency and Legislator Effects in Legislative Representation: Black Legislators or Black Districts? *Legislative Studies Quarterly*, *30*(3), 361–373.

Grömping, U. (2006). Relative Importance for Linear Regression in R: The Package relaimpo. *Journal of Statistical Software*, *17*(1), 1–27.

Levendusky, M. S. (2008). Clearer Cues, More Consistent Voters: A Benefit of Elite Polarization. *Political Behavior*, *30*(2), 167–190.

Mansbridge, J. (1999). Should Blacks Represent Blacks and Women Represent Women? A Contingent "Yes". *The Journal of Politics*, *61*(3), 628–657.

McCarty, N., Poole, K. T., & Rosenthal, H. (2001). Political Polarization and Income Inequality. *American Political Science Review*, *95*(2), 425–436.

McCarty, N., Poole, K. T., & Rosenthal, H. (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Pitkin, H. F. (1967). *The Concept of Representation*. Berkeley: University of California Press.

Poole, K. T., & Rosenthal, H. (1997). *Congress: A Political–Economic History of Roll Call Voting*. New York: Oxford University Press.

Shor, B., & McCarty, N. (2011). The Ideological Mapping of American Legislatures. *American Political Science Review*, *105*(3), 530–551.

Squire, P. (2012). *The Evolution of American Legislatures: Colonies, Territories, and States, 1619–2009*. Ann Arbor: University of Michigan Press.

Welch, S., & Bledsoe, T. (1985). Urban Reform and Its Consequences: A Study in Representation. *American Journal of Political Science*, 29(1), 52–80.

Features

FEATURES

The Hunt for Earth 2.0: An Interview with Professor Andrew Szentgyorgyi

Photo courtesy of NASA/JPL-Caltech

Ori Shi '29

Humanity has marveled at the mystery of the cosmos since the beginning of civilization. Although we first searched for answers with our naked eyes, we eventually developed telescopes for magnified observation, then invented complex electronic instruments to "see"—or rather, translate—numerous wavelengths beyond visible light. One leading pioneer of such technology today is Dr. Andrew Szentgyorgyi, Principal Investigator of the Giant Magellan Telescope (GMT)-Consortium Large Earth Finder (G-CLEF) project at the Harvard & Smithsonian Center for Astrophysics (CfA). Having previously worked across neutrino physics, high-energy astrophysics, and astronomical instrumentation, Dr. Szentgyorgyi is now building a spectrograph to detect Earth-like exoplanets.

Ori Shi (OS): Thank you for taking the time to speak with me! I would love it if you could tell me a little bit about yourself and what brought you to the Center for Astrophysics.

Andrew Szentgyorgyi (AS): I went to high school for three years, and I do not have a high school diploma. I tried very, very hard for my PhD, but I couldn't make my first degree happen. So I went to Syracuse University School of Art for a while. It was the Great Recession of 1971–1972, so eventually my father couldn't pay the tuition, and I was thrown out of art school. After a bunch of jobs, I got a job in a machine shop in a physics department. I decided physics was a lot easier than art, so I majored in physics. After I started doing particle physics and cosmic ray physics about 33 years ago, Dan Fabricant—who's down the hall—and I started a group building a number of robotic things for the MMT Observatory, which is operated by the University of Arizona and the Smithsonian Institution. Now, the group that we have today builds fairly large instrumentation for really big telescopes, such as the Giant Magellan Telescope (GMT).

The instrument I'm building is called the GMT-Consortium Large Earth Finder (G-CLEF). It's the spectrograph to search for Earth analogs—or Earth 2.0s—and then interrogate their atmospheres for biomarkers.

It turns out that to find Earth 2.0, you don't actually find the planet itself. You measure the motion of the star

that it orbits. As the planet goes in a big orbit, the sun itself shifts just a little tiny bit in response to the planet's gravity. And as the star gets pulled back and forth, you look at the shift of its light's detected velocity (called a velocity shift), like a policeman measuring the speed of a car with his radar gun. In fact, the maximum velocity shift is 10 centimeters a second. That's about as fast as a Galapagos turtle walks.

So, to separate the shifts due to the planet from Earth's environmental effects, we put the spectrograph in a vacuum chamber. It's like the high-tech version of what they make surfboards out of—like fiberglass—except it's a $3 million machine.

OS: Could you walk me through how exactly G-CLEF will use this light shift to detect oxygen?

AS: This is what the end product looks like: The G-CLEF breaks the light up into this high-resolution spectrum (Fig. 1). We optically take each little chunk and separate them out vertically, so it fits on a rectangle. Each one of these dark bands corresponds to something in the outer atmosphere of the star, the space between us and the star, or the Earth's atmosphere. And so you can use these bands to find out what the composition is.

But also, when you look at stars, you experience something called the Doppler shift. When the star moves towards you, this whole spectrum gets a little bluer. When it's moving away from you, the spectrum gets a little redder. And then, because our instrument is very stable, we can sense the motion of these dark lines going back and forth as the star moves, as I say, 10 centimeters a second.

This is the spectrum of a star called Procyon (the eighth brightest star in our night sky and part of the constellation Canis Minor). Here (orange boxes), you can see it has sodium in it. This big dark thing up here (green box) is because there's hydrogen. The important thing is, you see up here (blue box), there are these little, sort of regular dark bands. Those aren't the stars. They're here because light has to come through the Earth's atmosphere, and what you're seeing is the absorption of molecular diatomic oxygen in the Earth's atmosphere. That's really interesting, because the presence of oxygen is entirely due to life on Earth. Earth's atmosphere is about 21% diatomic oxygen, which doesn't happen naturally without life. If you turned off life on the surface of the Earth, that 21% oxygen would go down to a fraction of a percent in about 45,000 years. In terms of cosmic time, that's a heartbeat. That's why some of us feel strongly that the right place to search for evidence of life beyond the solar system is to look for these diatomic oxygen hallmarks in the spectra of exoplanets. When the exoplanet goes in front of the star, the star's light comes through the exoplanet's atmosphere, allowing us to see that atmosphere's makeup.

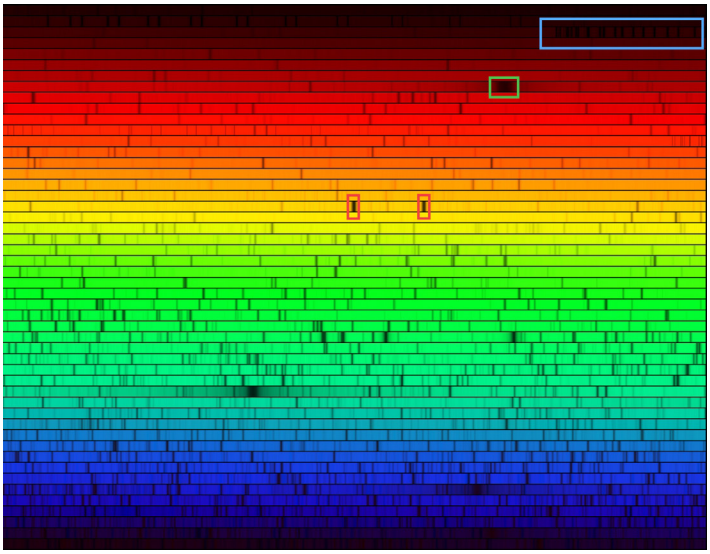


Figure 1. Light spectrum from the star Procyon featuring absorption lines from sodium (red boxes) and hydrogen (green box) in the star's gas, as well as Earth's atmospheric oxygen (blue box). Courtesy of Dr. Andrew Szentgorygi.

"If you turned off life on the surface of the Earth, that 21% oxygen would go down to a fraction of a percent in about 45,000 years. In terms of cosmic time, that's a heartbeat."

You might be worried that the oxygen in the exoplanet's spectra would be swamped by the Earth's atmosphere. But it turns out that because of the Doppler effect, all of those exoplanet atmosphere lines are shifted a little bit, so they're not right on top of Earth's lines.

To get these spectra, we collect light with the telescope and send it to the G-CLEF through an optical fiber. Inside the instrument, the light fans out, and ultimately, multiple cameras make an image right there on a charge-coupled device, the same thing in your iPhone.

OS: Since the G-CLEF is one component of the much larger GMT, could you tell me more about who else is involved in the GMT and the topics they're exploring?

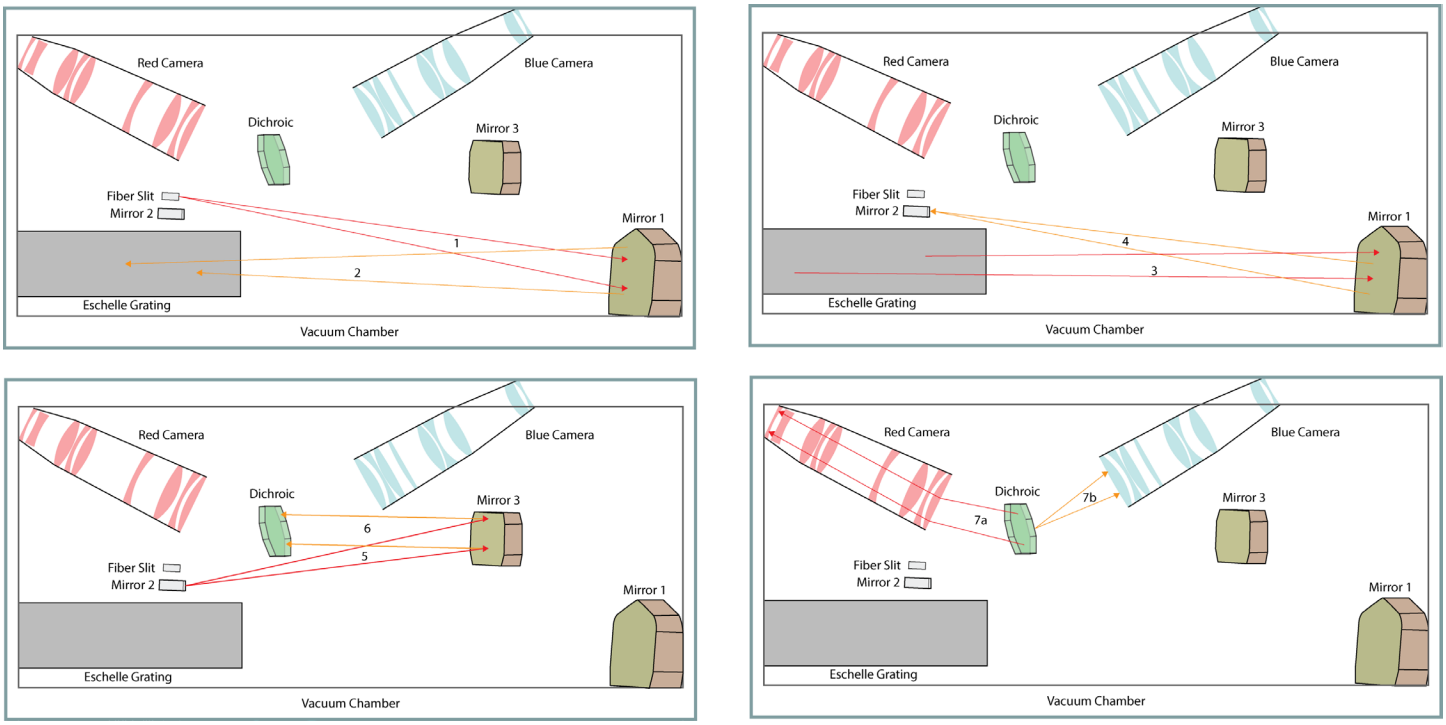


Figure 2. Cross-section of the G-CLEF showing the sequential path of received light (Steps 1–7). Courtesy of Iris Sung.

AS: The GMT is actually a large consortium. The CfA is leading the development, but it also involves people from the University of Chicago, Carnegie Mellon University, the Massachusetts Institute of Technology, the Taiwanese Institute of Astronomy, the Korea Astronomy and Space Sciences Institute, the Weizmann Institute of Science in Israel, the University of São Paulo in Brazil, and many more. We're also now expecting that the National Science Foundation will become a major partner, which means the whole United States astrophysics community will have access.

People will be doing a lot of different things. They'll be looking at supernovas and doing research about dark energy in the universe. They'll also be studying the earliest phases of structure formation in the universe and how planets and stars form. Other than Earth 2.0s, scientists will be using the G-CLEF to look for "ultra metal poor stars," which are the fossils of the earliest stars formed 13 to 14 billion years ago.

In the last 40 years, we realized that the universe isn't just a picture—it's a movie. It's a very, very dynamic place. So now, many people are studying what we call transient phenomena, which are stars exploding and black holes colliding. Many groups in the consortium will have a note that hangs on their boards that says, "If this kind of star goes off, stop doing what you're doing and look at it."

OS: We've talked a lot about the scientific importance of the G-CLEF and the GMT. But there are many who wonder if it's worthwhile to spend billions on instruments that look unreachably far away. How would you describe the value of instruments such as the GMT to society?

AS: There are lots of different takes on that. What I'm about to tell you is personal. I think that there's a tendency to try and find practical applications that we develop in astronomy. I used to keep a list of things that astronomers discovered. But the list of things that aren't curiosity-driven research is actually very small. And so my answer to you is that it's important to cure disease, develop technology, grow the economy, and all of those wonderful things that keep people well fed and safe. But if you don't interrogate what man's place in the universe is, all that stuff becomes essentially meaningless.

"In the last 40 years, we realized that the universe isn't just a picture—it's a movie."

And I don't just mean man's place in the universe in terms of whether there's life elsewhere. The universe is investigated by everybody: the humanities, philosophy,

and creative arts. These are all endeavors to supply context for what it means to be human, and the minute you stop asking that question, you're just an animal.

OS: That's very insightful. Speaking of life elsewhere, do you think it is probable that the G-CLEF will detect diatomic oxygen?

AS: Oh, well, I'm very agnostic about that. I don't want to pretend for a minute that I'm some gleaming hero who is finally going to answer one of the most fundamental questions. On the other hand, it's probably the first time in history where we really can, in a qualitative way, answer the question of how likely it is that we're alone in the universe. Even a null result would be really, really huge. I don't really care, almost. I'm mostly interested in finding out more.

I do want to say that, if I find evidence of diatomic oxygen in nearby exoplanets, that's not the end of the story. The European Southern Observatory is building a similar instrument. And if we find O_2 first, we need them to find it too and confirm. People always say, "Well, how do you know you *need* water and oxygen for life?" The answer is, I don't. But that's why we have to do this search.

This interview has been edited for length and clarity.

When Does Care Overreach? The Harms of Overmedicalization in Mental Health

Image courtesy of PxHere

Triscia Afihene '27

Introduction

In 2024, 61.5 million adults in the United States reported a diagnosed mental health disorder, a figure that only continues to grow (National Alliance on Mental Illness [NAMI], 2025). Fortunately, this trend has been accompanied by increasing mental health awareness among not only scholars and clinicians but also the public. In the 1980s and 1990s, programs such as NAMI's destigmatization campaigns and Congress's "Decade of the Brain" initiative began to recognize the importance of mental health research and mainstream education (Walthall, 2020). Since then, they have successfully encouraged Americans to view mental health as a vital aspect of overall well-being. Through public service announcements, celebrity advocacy, and educational curricula, efforts to address stigma have validated individual experiences and connected millions with life-changing treatment, establishing mental illness as a legitimate health concern that warrants attention, compassion, and care (Walthall, 2020; Waqas et al., 2020). While it is difficult to establish a direct causal relationship, many scholars attribute the expansion of mental health diagnoses to the rise in mental health literacy via awareness efforts (Haslam & Tse, 2024).

Simultaneously, however, the growing visibility of mental health has subtly shifted the boundaries of what is considered "pathological." Conditions once considered natural emotional fluctuations, such as low sexual arousal and even intense grief following the death of a loved one, began to fall within the bounds of psychiatric classification (American Psychiatric Association, n.d.; Eske, 2020). Another example is premenstrual dysphoric disorder (PMDD), which consists of symptoms such as decreased interest in normal activities, breast swelling or tenderness, and mood swings beginning during the week before menstruation (Johns Hopkins Medicine, n.d.). PMDD became a mental health condition in the Fifth Edition of the American Psychological Association's (APA) *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5)—but both psychologists and feminist scholars have criticized the addition of PMDD in DSM-5, claiming that it leads to overdiagnosis of normal hormonal changes or the disempowering of women by claiming women are pathologically emotional, reinforcing the "myth of

the irrational woman" (Schroll & Lauritsen, 2022; King, 2020).

Thus, while the medicalization of mental health has helped many people better understand and manage their challenges, it also carries risks. The evolution of the DSM from the Third to Fifth Edition has resulted in notable diagnostic inflation: 51 disorders displayed net inflation (Fabiano & Haslam, 2020). The expanding reach of diagnostic criteria may expose individuals to unnecessary medical interventions, pharmaceutical side effects, and a narrowing of how society understands human emotion and behavior. Ultimately, the motivation to identify, name, and treat—though rooted in good intentions—can, when extended beyond clinical significance, distort medicine's healing power into the source of harm itself.

Pathologizing Pain: The DSM and the Cost of Medicalization

The modern tendency to medicalize human hardship can be traced back to the evolution of the DSM, which gradually transformed psychiatry from a loosely interpretive, psychoanalytic discipline into one grounded in diagnostic specificity. The publication of the Third Edition of the manual, the DSM-III, in 1980 marked a turning point (Shorter, 2015). Moving away from the individualized diagnoses of psychoanalysis, the DSM-III adopted explicit, consensus-based criteria that imbued psychiatry with a new sense of legitimacy and precision. For many, this shift was long overdue, and allowed many patients to finally receive standardized diagnostic labels and care, validating their experiences.

"These findings raise an uneasy paradox: as treatment efforts expand, so too does the population living with long-term iatrogenic mental health challenges."

Yet, the expansion of diagnostic criteria came with a cost. The DSM-III's standardized framework helped clinicians communicate more clearly and enabled researchers to study mental illness systematically, but it also laid the groundwork for a growing, potentially exploitative pharmaceutical market (Nikkel & Whitaker, 2018). Following the publication of the DSM-III, billions of dollars were invested in psychopharmacological research, leading to the development of drugs like Prozac, the first selective serotonin reuptake inhibitor (SSRI; Kawa & Giordano, 2012).

While these developments dramatically improved the lives of many by providing effective treatment options, they also reinforced the idea that emotional suffering could be treated simply through biochemical correction (Shpancer, 2022). Furthermore, psychiatric medications such as antidepressants carry risks and limitations. Studies and several analyses suggest that some medications may offer short-term improvement compared to a placebo, but can also increase the likelihood of relapse or chronic symptoms over time by 36.4% (Nikkel & Whitaker, 2018; Kirsh, 2019; Batelaan et al., 2017). These findings raise an uneasy paradox: as treatment efforts expand, so too does the population living with long-term iatrogenic mental health challenges—conditions that arise from unintended consequences of medical treatment itself (Moynihan, 2012). Despite the DSM-III's success in legitimizing care, its legacy invites reflection on how a system built to heal can expose individuals to unnecessary medicalization and potential harm.

The Hidden Harms of Overdiagnosis

As diagnostic categories multiplied and pharmaceutical solutions expanded, the threshold for what qualified as a disorder also became increasingly expansive, with criteria for some disorders becoming relaxed (Fabiano & Haslam, 2020). In a study done on a Baltimore population, it was found that of current antidepressant users 38% did not meet the requirements for OCD, major depressive disorder (MDD), or panic disorder, and 69% did not meet the requirements for MDD specifically (Takayanagi et al, 2015). Conversely, research has shown that antidepressants achieve almost no benefit compared to a placebo in mild and moderate depression (Mayor, 2008). What began as a movement toward inclusion and understanding slowly revealed another face, one where overdiagnosis and overmedication exposed individuals to risks that the system was never designed to anticipate.

Since the 1970s, several scholars—including philosopher Ivan Illich and, more recently, influential psychiatrist Allen Frances, who served as Chair of the DSM-IV Task Force—have argued that medicine's good intentions can inadvertently cause harm to patients by transforming a broad range of experiences into diagnosable problems that require treatment. As medicine extended its reach, many pharmaceutical interventions once reserved for severe illness are now used for mild or transient distress, resulting in often quite significant

chronic iatrogenic illness (Moynihan et al., 2002). An example can be seen in the expanded use of Olanzapine (Zyprexa), an antipsychotic traditionally prescribed to treat schizophrenia and bipolar disorder, for insomnia (Khaledi-Paveh et al., 2021). Psychiatric medications, such as antidepressants, while life-saving for many, can lead to significant side effects such as the inability to concentrate, facial rigidity, a distorted sense of self, and withdrawal-like symptoms that linger even after the treatments stop. Some patients on psychiatric medications have described feeling "numb" or like a "zombie" to illustrate the handicap that they become over time (Rodriguez Del Barrio et al., 2014). Antidepressant discontinuation syndrome, for example, is a growing issue, with studies showing that approximately 20% of patients experience withdrawal symptoms after discontinuing antidepressant medications after at least six weeks encouraging them to go back on antidepressants (Warner et al., 2006).

Nevertheless, it is important to acknowledge that overdiagnosis often stems from compassion rather than negligence. Physicians are trained to quickly identify straightforward causes of and treatments for pain; in a healthcare environment where time is often limited, prescribing medication can appear to be the most efficient route to care (Kale & Korenstein, 2018; Wakefield, 2015). Patients, too, often seek explanations for distress that feels unmanageable, turning to medicine for clarity and a way to validate their experience. Yet this shared "risk aversion" creates a loop where the fear of under-treating illness outweighs the caution against unnecessary, and potentially harmful intervention (Hofmann, 2016). Structural forces surrounding psychiatry reinforce this pattern: for instance, pharmaceutical companies, driven by profit and enabled by expanding diagnostic criteria, market new medications directly to both physicians and the public, framing natural variations in mood and behavior as treatable disorders. An example can be seen in biopharmaceutical giant GSK's marketing for the antidepressant Paxil, a drug targeting social anxiety disorder (Wolinski, 2005). These tactics foster a medical culture where care equates to pharmacological treatment, and the side effects that promote long-term use make the maintenance of the disorder more unbearable than the illness (Rodriguez Del Barrio et al., 2014). When care becomes synonymous with medication, we risk treating discomfort as disease and exposure to treatment as healing, even when it quietly deepens suffering. A shift towards cognitive therapies—grounded in self-reflection and behavioral change rather than biochemical correction—offers a path to healing that offers relief without the risks of medication (DeRubeis et al., 2009).

"When care becomes synonymous with medication, we risk treating discomfort as disease and exposure to treatment as healing, even when it quietly deepens suffering."

Medicalization Today: Gen Z and the Normalization of Diagnosis

The expansion of psychiatric care did not remain confined to clinics or hospitals—it gradually entered society itself. The push to encourage acceptance and bring validity to mental health and disorders fostered a powerful relationship between pathologization and the media. The efforts of NAMI to bridge psychiatry and the public marked a critical turning point: mental health was no longer a private issue but a collective concern discussed openly. As understanding deepened and awareness spread, the language of mental health became increasingly familiar.

Campaigns like the 2017 World Health Organization's (WHO) "Depression: Let's Talk" encouraged individuals to notice and name negative psychological states, an important step toward help-seeking, but one that may also lead some to equate psychological states such as temporary sadness with clinical depression (Foulkes & Andrews, 2023). Individuals, often because they believe that the experiences they are having are not synonymous with those of a small group of peers, become convinced that their "symptoms" are indicative of a mental disorder, even when they are common experiences among the general public (Suhr & Johnson, 2022).

The COVID-19 pandemic further intensified this dynamic with increased use of social media. With isolation confining many, particularly Gen Z, to digital spaces, social media became both a source of community and a new space for mental health discourse (Foulkes & Andrews, 2023). In fact, 25% of young people have used social media to self-diagnose, with Gen Z constituting a large proportion (Redshaw, 2023). According to the WHO (2022), during the pandemic, there was a 25% increase in reports of anxiety and depression; in the absence of accessible healthcare, digital spaces often fill the gap, encouraging self-diagnosis and self-treatment.

Research has shown that during and after the pandemic, rates of psychotropic medication prescribing increased among patients with depressive disorders (Luo et al., 2024). While this openness has helped normalize conversations around mental health, it also perpetuates a culture in which emotional struggles are immediately medicalized—where understanding the self increasingly begins with searching for symptoms.

Conclusion

The expansion of clinical psychiatry has sought to bring understanding and compassion to those experiencing mental illness, a mission that has given voice to countless forms of suffering once silenced or dismissed. Yet the same expansion that made care accessible has also fueled the temptation to view all forms of distress through a medical lens. When every discomfort becomes a symptom to treat, medicine risks drifting away from its original purpose: to heal where healing is needed and to help people live meaningfully with what cannot be cured. The problem is not medicine's desire to alleviate pain and suffering, but that it risks turning care into exposure, introducing individuals to medications and side effects that they could have gone without (Cassel, 1982). In a study done on a Baltimore population, it was found that of current antidepressant users 38% did not meet the requirements for OCD, major depressive disorder (MDD), or panic disorder, and 69% did not meet the requirements for MDD specifically (Takayanagi et al., 2016). Conversely, research has shown that antidepressants achieve almost no benefit compared to a placebo in mild and moderate depression (Mayor, 2008). In the effort to ensure that no illness goes untreated, medicine sometimes forgets that restraint is also a form of care. Prescribing a drug is an act of trust, and when used without necessity, can harm the very body it was meant to heal and protect (Kale & Korenstein, 2018). The consequences invite a difficult question: how can medicine balance its desire to help with its Hippocratic obligation to "do no harm"?

References

American Psychiatric Association. (n.d.). *Prolonged grief disorder*. Retrieved November 22, 2025, from https://www.psychiatry.org:443/patients-families/prolonged-grief-disorder

Ballard, K., & Elston, M. A. (2005). Medicalisation: A multi-dimensional concept. *Social Theory & Health*, *3*(3), 228–241. https://doi.org/10.1057/palgrave.sth.8700053

Batelaan, N. M., Bosman, R. C., Muntingh, A., Scholten, W. D., Huijbregts, K. M., & van Balkom, A. J. L. M. (2017). Risk of relapse after antidepressant discontinuation in anxiety disorders, obsessive-compulsive disorder, and post-traumatic stress disorder: Systematic review and meta-analysis of relapse prevention trials. *The BMJ*, *358*, j3927. https://doi.org/10.1136/bmj.j3927

Cassel, E. J. (1982). The nature of suffering and the goals of medicine. *New England Journal of Medicine*, *306*(11), 639–645. https://doi.org/10.1056/NEJM198203183061104

Davidson, M., & Gabos-Grecu, C. (2020). Do DSM classifications help or hinder drug development? *Dialogues in Clinical Neuroscience, 22*(1), 73–79. https://doi.org/10.31887/DCNS.2020.22.1/mdavidson

DeRubeis, R. J., Siegle, G. J., & Hollon, S. D. (2008). Cognitive therapy vs. medications for depression: Treatment outcomes and neural mechanisms. *Nature Reviews Neuroscience*, *9*(10), 788–796. https://doi.org/10.1038/nrn2345

Eske, J. (2020, November 26). *What is female arousal disorder? Symptoms and treatment*. Medical News Today. https://www.medicalnewstoday.com/articles/female-arousal-disorder

Fabiano, F., & Haslam, N. (2020). Diagnostic inflation in the DSM: A meta-analysis of changes in the stringency of psychiatric diagnosis from DSM-III to DSM-5. *Clinical Psychology Review*, *80*, 101889. https://doi.org/10.1016/j.cpr.2020.101889

Foulkes, L., & Andrews, J. L. (2023). Are mental health awareness efforts contributing to the rise in reported mental health problems? A call to test the prevalence inflation hypothesis. *New Ideas in Psychology*, *69*, 101010. https://doi.org/10.1016/j.newideapsych.2023.101010

Haslam, N., & Tse, J. S. (2025). Public awareness of mental illness: Mental health literacy or concept creep? *Australasian Psychiatry*, *33*(1), 18–20. https://doi.org/10.1177/10398562241292202

Hofmann, B. (2016). Medicalization and overdiagnosis: Different but alike. *Medicine, Health Care and Philosophy*, *19*(2), 253–264. https://doi.org/10.1007/s11019-016-9693-6

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, *36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1

Johns Hopkins Medicine. (n.d.). *Premenstrual dysphoric disorder (PMDD)*. https://www.hopkinsmedicine.org/health/conditions-and-diseases/premenstrual-dysphoric-disorder-pmdd

Kale, M. S., & Korenstein, D. (2018). Overdiagnosis in primary care: Framing the problem and finding solutions. *BMJ*, k2820. https://doi.org/10.1136/bmj.k2820

Kawa, S., & Giordano, J. (2012). A brief historicity of the Diagnostic and Statistical Manual of Mental Disorders: Issues and implications for the future of psychiatric canon and practice. *Philosophy, Ethics, and Humanities in Medicine*, *7*(1), 2. https://doi.org/10.1186/1747-5341-7-2

Khaledi-Paveh, B., Maazinezhad, S., Rezaie, L., & Khazaie, H. (2021). Treatment of chronic insomnia with atypical antipsychotics: Results from a follow-up study. *Sleep Science*, *14*(1), 27–32. https://doi.org/10.5935/1984-0063.20190149

King, S. (2020). Premenstrual syndrome (PMS) and the myth of the irrational female. In C. Bobel, I. T. Winkler, B. Fahs, K. A. Hasson, E. A. Kissling, & T.-A. Roberts (Eds.), *The Palgrave Handbook of Critical Menstruation Studies* (pp. 287–302). Springer Singapore. https://doi.org/10.1007/978-981-15-0614-7_23

Kirsch, I. (2019). Placebo effect in the treatment of depression and anxiety. *Frontiers in Psychiatry*, *10*, 407. https://doi.org/10.3389/fpsyt.2019.00407

Luo, H., Chai, Y., Li, S., Lau, W. C. Y., Torre, C. O., Hayes, J., Lam, I. C. H., Lin, X., Yin, C., Fortin, S., Kern, D. M., Lee, D. Y., Park, R. W., Jang, J.-W., Chui, C. S. L., Li, J., Seager, S., Man, K. K. C., & Wong, I. C. K. (2024). Psychotropic drug prescribing before and during the COVID-19 pandemic among people with depressive and anxiety

disorders: A multinational network study. *The Lancet Psychiatry*, *11*(10), 807–817. https://doi.org/10.1016/S2215-0366(24)00245-1

Mayor, S. (2008). Meta-analysis shows difference between antidepressants and placebo is only significant in severe depression. *BMJ*, *336*(7642), 466.1-466. https://doi.org/10.1136/bmj.39503.656852.DB

Moynihan, R. (2012). *Disease-mongering: Widening the boundaries that define medical illness*. Health Action International. https://haiweb.org/encyclopaedia/disease-mongering/

National Alliance on Mental Illness. (n.d.). *Mental health by the numbers*. Retrieved November 22, 2025, from https://www.nami.org/about-mental-illness/mental-health-by-the-numbers/

Nikkel, R., & Whitaker, R. (2018, October 22). Flooding the world with psychiatric drugs could boost the burden of mental disorders. *STAT*. https://www.statnews.com/2018/10/22/flooding-world-psychiatric-drugs-boost-burden-mental-disorders/

Redshaw, L. (2023, August 2). Is self-diagnosis on social media helping or hurting people's health? *Tebra*. https://www.tebra.com/theintake/medical-deep-dives/tips-and-trends/is-self-diagnosis-on-social-media-helping-or-hurting-peoples-health

Robinson, J. (2022, August 31). The serotonin theory of depression: How the media got it all wrong. *The Pharmaceutical Journal*. https://pharmaceutical-journal.com/article/feature/the-serotonin-theory-of-depression-how-the-media-got-it-all-wrong

Rodriguez Del Barrio, L., Onocko Campos, R., Stefanello, S., Vianna Dantas Dos Santos, D., Cyr, C., Benisty, L., & De Carvalho Otanari, T. (2014). Human rights and the use of psychiatric medication. *Journal of Public Mental Health*, *13*(4), 179–188. https://doi.org/10.1108/JPMH-06-2013-0039

Schroll, J. B., & Lauritsen, M. P. (2022). Premenstrual dysphoric disorder: A controversial new diagnosis. *Acta Obstetricia et Gynecologica Scandinavica*, *101*(5), 482–483. https://doi.org/10.1111/aogs.14360

Shorter, E. (2015). The history of nosology and the rise of the Diagnostic and Statistical Manual of Mental Disorders. *Dialogues in Clinical Neuroscience*, *17*(1), 59–67. https://doi.org/10.31887/DCNS.2015.17.1/eshorter

Shpancer, N. (2022, July 24). Depression is not caused by chemical imbalance in the brain. *Psychology Today*. https://www.psychologytoday.com/us/blog/insight-therapy/202207/depression-is-not-caused-chemical-imbalance-in-the-brain

Suhr, J. A., & Johnson, E. E. H. (2022). First do no harm: Ethical issues in pathologizing normal variations in behavior and functioning. *Psychological Injury and Law*, *15*(3), 253–267. https://doi.org/10.1007/s12207-022-09455-z

Takayanagi, Y., Spira, A. P., Bienvenu, O. J., Hock, R. S., Carras, M. C., Eaton, W. W., & Mojtabai, R. (2015). Antidepressant use and lifetime history of mental disorders in a community sample: Results from the Baltimore Epidemiologic Catchment Area Study. *The Journal of Clinical Psychiatry*, *76*(01), 40–44. https://doi.org/10.4088/JCP.13m08824

Walthall, J. (2020, June 1). The evolution of the mental health movement. *National Alliance on Mental Illness*. https://www.nami.org/Blogs/NAMI-Blog/June-2020/The-Evolution-of-the-Mental-Health-Movement

Warner, C. H., Bobo, W., Warner, C., Reid, S., & Rachal, J. (2006). Antidepressant discontinuation syndrome. *American Family Physician*, *74*(3), 449–456.

Waqas, A., Malik, S., Fida, A., Abbas, N., Mian, N., Miryala, S., Amray, A. N., Shah, Z., & Naveed, S. (2020). Interventions to reduce stigma related to mental illnesses in educational institutes: A systematic review. *The Psychiatric Quarterly*, *91*(3), 887–903. https://doi.org/10.1007/s11126-020-09751-4

Wolinsky, H. (2005). Disease mongering and drug marketing. *EMBO Reports*, *6*(7), 612–614. https://doi.org/10.1038/sj.embor.7400476

World Health Organization. (2022, March 2). *COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide*. https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide

Galactic Gold Rush: The Promise and Perils of Mining Beyond Earth

Image courtesy of NASA, 2018

Devi Kuscer ’29

Introduction

For decades, humanity has looked beyond Earth for solutions to terrestrial problems. As global warming and resource depletion intensify, worlds outside our own are viewed not as an unrealistic hope, but as a practical solution for humanity’s future survival. Reserves of critical elements, such as platinum, are in short supply, causing costs to rapidly rise and directly endanger continued technological development (Andrews et al., 2015). However, the mining of these commodities often involves unethical and unsustainable practices (Iles, 2020). Space mining—often called extraterrestrial resource extraction (ERE)—has the potential to alleviate the Earth of these problems through harvesting materials like water, gold, and platinum from extraterrestrial bodies.

However, this ambitious project is accompanied by a steep ethical cost. Pristine extraterrestrial environments could become polluted, and nations or corporations with the greatest wealth could monopolize the industry, leading to an even larger wealth gap. Given the long history of global exploitation—such as the destruction of around 20% of the Amazon rainforest—is it possible for ERE to avoid propagating humanity’s terrestrial mistakes beyond Earth?

The Feasibility and Current Targets of ERE

Due to the high costs of launching space missions, space mining remains a purely theoretical endeavor. For example, NASA’s recent OSIRIS-Rex mission to the asteroid Bennu, which collected and returned a sample of the asteroid’s surface, cost around $1.16 billion (Cost of OSIRIS-REx, 2020). However, the increase in the amount of material collected from celestial bodies signifies major developments in ERE technology. Nevertheless, these missions have yet to return with samples large enough to make space mining commercially viable.

A difficulty with making space mining feasible is the high upfront cost stemming from the development

and launch of custom-built spacecraft, with one model estimating the total cost of a mission to the main asteroid belt at around $100 billion (Probst et al., 2019). This problem could be addressed through innovative approaches aiming to increase the rate of material return to Earth, such as the utilization of smaller, mass-produced spacecraft, which would both increase the processing rate of mined materials and reduce the development costs of missions (Calla et al., 2018).

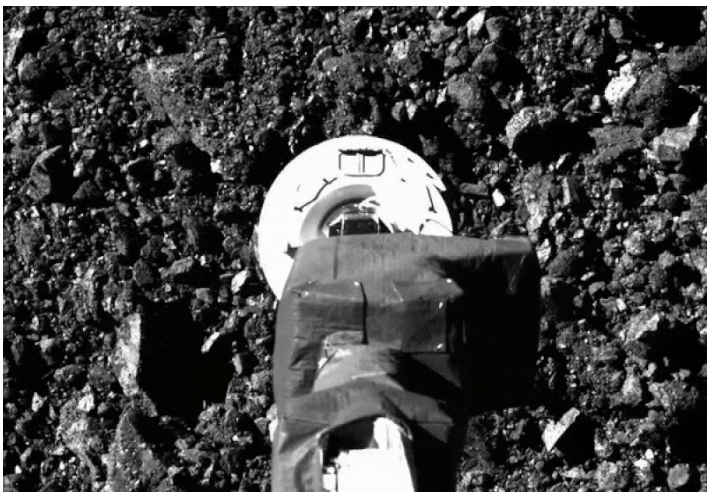


Figure 1. OSIRIS-Rex collecting a sample of the surface of Bennu (NASA/Goddard/University of Arizona, 2020)

Additionally, the chosen resource and its intended market destination play a large role in the terrestrial success of ERE. For asteroid-mined platinum group metals (PGMs) to become practical, their quantity must be substantial enough to substitute an equivalent amount of terrestrial material—over 100,000 tons of platinum. Thus, at present, the most economically viable path appears to be mining volatiles—such as water and ammonia— for astronaut use at mission destinations instead of transporting supplies from Earth, a process referred to as In Situ Resource Utilization (ISRU) (NASA, 2023a).

ISRU is a principal objective of NASA's Artemis III mission, which is scheduled to return people to the moon by 2027. By sustaining the mission's crew through ISRU, space exploration would become far more efficient and cost-effective (NASA, 2023b). The development of technology used for this purpose also marks a crucial step in ERE, as it could potentially be adapted for future metal-focused extraction.

A major milestone in ERE came in 2023, when the OSIRIS-Rex spacecraft collected and returned a 121.6 gram sample of Bennu's carbon-rich surface—the largest amount of space material brought to Earth to date (Gaskill, 2025). This record-breaking sample demonstrates the potential for the production of goods with space-mined resources. For example, catalytic converters, a crucial component of vehicles which convert toxic exhaust fumes into less harmful substances, require just 3–7 grams of platinum per unit, meaning that even the small sample from Bennu could supply multiple converters (APMEX, 2023). By adapting mission design to return larger quantities of material, the industry is demonstrating clear progress towards a scalable future.

Potential candidates for the mining of Earth-rare metals have already been identified. Asteroid 1986DA is four times closer to the Earth than Bennu, and has been estimated to contain nickel, iron, cobalt, and platinum in amounts exceeding terrestrial supplies. If Asteroid 1986DA were mined and the extracted metals sold over 50 years, they could potentially be worth around $11.65 trillion (Sanchez et al., 2021). This high profit showcases the allure of the industry despite current setbacks.

Current and Potential Legal and Governmental Challenges Regarding ERE

The pursuit of space mining is currently complicated by a significant legal vacuum and increasing geopolitical conflict over ownership. The foundational 1967 Outer Space Treaty (OST) established a legally gray area: although it strictly forbids claims of sovereignty over celestial bodies by any single nation, it does not discuss the exploitation of extraterrestrial resources by private companies (UNOOSA, 1966). As a result, nations could take advantage of this ambiguity by using private companies as a legal bypass to indirectly establish extra-terrestrial monopolies. Although newer frameworks have attempted to establish principles that account for modern problems, they come with their own limitations.

Some countries have already begun to adapt to the age of the commercial exploitation of space. The U.S. (US

Figure 2. Animation of Artemis explorers on the Moon (NASA, 2025)

Congress, 2015) and Japan (Japanese Law Translation, 2021) have passed domestic laws that grant their national companies legal rights over the resources they extract in space. This action essentially bypassed the communal spirit of the OST by granting rights to space resources to private companies rather than the nation, thereby creating a dangerous legal precedent where access to and ownership of valuable resources like water and precious metals are determined by national jurisdiction rather than international consensus. Moreover, the potential for individual profit is equally large: American astrophysicist and public educator Neil deGrasse Tyson has predicted that "the first trillionaire there will ever be is the person who exploits the natural resources on asteroids" (Kramer, 2015). Tyson's statement contrasts the values of the OST, which presented space as "a common heritage of mankind," and further highlights the concern with how mined resources should be divided.

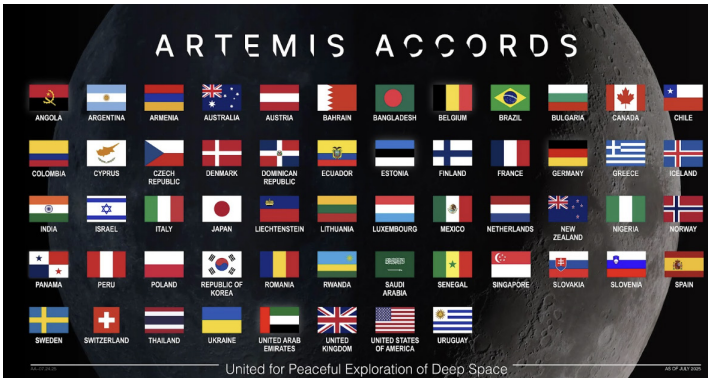


Figure 3. Signatory countries of the Artemis Accords (NASA, 2025)

The Artemis Accords, a non-binding agreement established in 2020, attempts to alleviate these concerns by providing a set of principles to further govern humanity's behavior in space (NASA, 2020). The framework, signed by 56 nations, explicitly states that resource extraction does not constitute a nation claiming ownership of a part of outer space, therefore ensuring that ERE does not go against the laws set by the OST. The accords also introduce the concept of "safety zones"—areas where signatories must provide notification of their activities and commit to coordination to avoid harmful interference. Such zones would allow for increased transparency and would lessen the possibility of conflict. However, as the framework is not legally binding, nations are not obligated to follow through on their commitments. Moreover, nations with major stakes in the space industry, such as Russia and China, have not signed the Accords, arguing that the agreement is too "US-centric" (Gross, 2023). Russia, in particular, has criticized the agreement's stance on ERE,

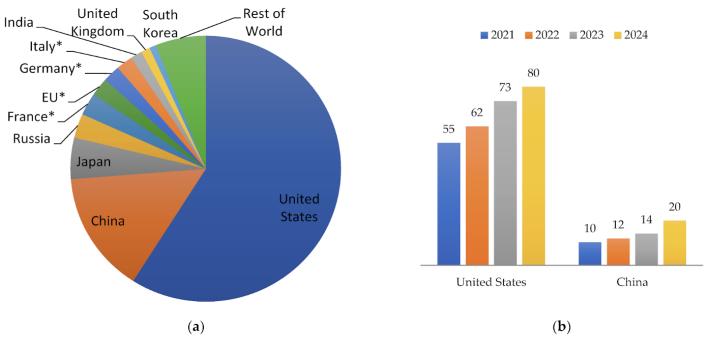


Figure 4. Government spending on space programs in billions of dollars. (a) global distribution, (b) comparison of US and China (Pietrzak, 2025)

comparing it to a form of colonialism. Sergei Savelyev, deputy head of international cooperation at Roscosmos, stated: "There have already been examples in history when one country decided to start seizing territories in its own interests and everyone remembers how that turned out" (Times, 2020). Such remarks beg the question of whether the accords will genuinely limit the monopolization of space resources, or whether they will encourage it.

Ethical, Environmental, and Equity Issues

Although space mining offers a solution to the pollution of the Earth's environment, it also has the potential to cause irreversible harm to extraterrestrial environments. Unregulated mining operations outside the Artemis Accords and the OST could introduce terrestrial contamination and create significant debris, destroying pristine environments that could be highly informative for scientific research on our solar system (Nelken-Zitser, 2025).

Another major ethical dilemma centers around who will benefit from the vast wealth of space. Unchecked resource utilization risks mirroring historical colonial exploitation, where powerful nations or wealthy corporations gain riches while developing countries are excluded (Rivkin et al., 2021). Nations with successful space programs—such as the US, Japan, and China—as well as private companies involved in space exploration have a clear advantage in ERE over developing countries without an established space sector, such as Bangladesh or Rwanda (Central Intelligence Agency, 2023). The concentration of spacefaring technology among only a few nations and corporations means that, by default, these entities would be the first to establish mining claims, further inhibiting the ability of developing nations to establish themselves in the global space industry (Kirby, 2024). Without robust, globally mandated safeguards, such as an unbiased global

authority that oversees extraterrestrial activities, global inequality could widen as space becomes a new source of wealth for the privileged few rather than the common heritage of all mankind.

Conclusion

Space mining offers the potential for the expansion of human development both on and beyond Earth. It promises a sustainable and highly profitable source of vital materials and a way to cut the costs of exploration; however, the question of whether humanity's destructive behaviour is eliminated completely or simply shifted to the realm of outer space remains unanswered.

Despite its exciting implications, commercial-scale extraction likely remains decades away. Costs are too high, current technology is underdeveloped, and humanity has yet to even return to the moon. However, real progress is constantly being made. NASA's Artemis III mission aims to drastically increase humanity's presence in space, and with the depletion of Earth's critical resources, the space-mining industry is expected to reach more than $23 billion by 2040 (Mordor Intelligence, 2025). These projections highlight humanity's determination to continue exploring the cosmos, expanding our reach far beyond Earth, and establishing ourselves as a true space-faring species.

References

Andreas Makoto Hein, Robert Matheson, Dan Fries. A Techno-Economic Analysis of Asteroid Mining. International Astronautical Congress 2018, Oct 2018, Bremen, Germany. ffhal-01910095f.

Andrews, D. G., Bonner, K. D., Butterworth, A. W., Calvert, H. R., Dagang, B. R. H., Dimond, K. J., Eckenroth, L. G., Erickson, J. M., Gilbertson, B. A., Gompertz, N. R., Igbinosun, O. J., Ip, T. J., Khan, B. H., Marquez, S. L., Neilson, N. M., Parker, C. O., Ransom, E. H., Reeve, B. W., Robinson, T. L., & Rogers, M. (2015a). Defining a successful commercial asteroid mining program. *Acta Astronautica*, 108, 106–118. https://doi.org/10.1016/j.actaastro.2014.10.034.

APMEX. (2023, August 16). *How Much Platinum is in a Catalytic Converter?* - APMEX. APMEX. https://learn.apmex.com/answers/how-much-platinum-is-in-a-catalytic-converter/.

Calla, P., Fries, D., & Welch, C. (2018). *Asteroid mining with small spacecraft and its economic feasibility.* https://doi.org/10.48550/arxiv.1808.05099.

Calla, P., Fries, D., & Welch, C. (2019). Asteroid mining with small spacecraft and its economic feasibility. *ArXiv:1808.05099 [Astro-Ph]*. https://arxiv.org/abs/1808.05099.

Central Intelligence Agency. (2023). *Space agency/agencies* . Cia.gov. https://www.cia.gov/the-world-factbook/about/archives/2023/field/space-agency-agencies

Cost of OSIRIS-REx. (2020, August 7). The Planetary Society. https://www.planetary.org/space-policy/cost-of-osiris-rex.

Gaskill, M. (2025, August 22). *NASA's Bennu Samples Reveal Complex Origins, Dramatic Transformation.* NASA Science. https://science.nasa.gov/missions/osiris-rex/nasas-bennu-samples-reveal-complex-origins-dramatic-transformation/.

Gross, M. (2023, January 27). *The Artemis Accords: International Cooperation in the Era of Space Exploration.* Harvard International Review. https://hir.harvard.edu/the-artemis-accords/.

Iles, A. M., Alastair. (2020). *HYLE 26-1 (2020): The Ethics of Rare Earth Elements Over Time and Space.* Www.hyle.org. https://www.hyle.org/journal/issues/26-1/martin.htm,

Japanese Law Translation. (2021, December 23). *Act on the Promotion of Business Activities for the Exploration and Development of Space Resources - English - Japanese Law Translation.* Www.japaneselawtranslation.go.jp. https://www.japaneselawtranslation.go.jp/en/laws/view/4332/en.

Kirby, H. (2024). *Spring 2024 Comment International Space Law in a New Space Age: What Laws Will Regulate Space Mining and NASA's Gateway Program? International Space Law in a New Space Age: What Laws Will Regulate Space Mining and NASA's Gateway Program?* https://www.smu.edu/-/.media/site/law/students/law-journals/spring-2024-comments/kirby_final.pdf.

Kramer, K. (2015, May 3). *Neil deGrasse Tyson Says Space Ventures Will Spawn First Trillionaire.* NBC News. https://www.nbcnews.com/science/space/neil-degrasse-tyson-says-space-ventures-will-spawn-first-trillionaire-n352271.

Mordor Intelligence. (2025, November 7). *Space Mining Market Report | Industry Analysis, Size & Forecast Overview.* Mordor Intelligence. https://www.mordorintelligence.com/industry-reports/space-mining-market-industry.

NASA. (2020). *THE ARTEMIS ACCORDS PRINCIPLES FOR COOPERATION IN THE CIVIL EXPLORATION AND USE OF THE MOON, MARS, COMETS, AND ASTEROIDS FOR PEACEFUL PURPOSES.* https://www.nasa.gov/wp-content/uploads/2022/11/Artemis-Accords-signed-13Oct2020.pdf?emrc=68f31bd19f945.

NASA. (2023a, July 26). *Overview: In-Situ Resource Utilization - NASA*. NASA. https://www.nasa.gov/overview-in-situ-resource-utilization/.

NASA. (2023b, September 30). *In-Situ Resource Utilization (ISRU) - NASA*. Nasa.gov. https://www.nasa.gov/mission/in-situ-resource-utilization-isru/.

Nelken-Zitser, J. (2025, February 25). *Moon mining could net billions, but astronomers say it's bad for science.* Business Insider. https://www.businessinsider.com/mining-moon-minerals-lucrative-bad-news-science-research-astronomists-2025-2.

Pietrzak, M. (2025). Economic and Social Aspects of the Space Sector Development Based on the Modified Structure–Conduct–Performance Framework. *World, 6(2), 79.* https://doi.org/10.3390/world6020079

Probst, A., Nitzl, C., Kraus, F., & Förstner, R. (2019). Cost Estimation of an Asteroid Mining Mission using Partial Least Squares Structural Equation Modelling (PLS-SEM). *Acta Astronautica, 167.* https://doi.org/10.1016/j.actaastro.2019.07.032.

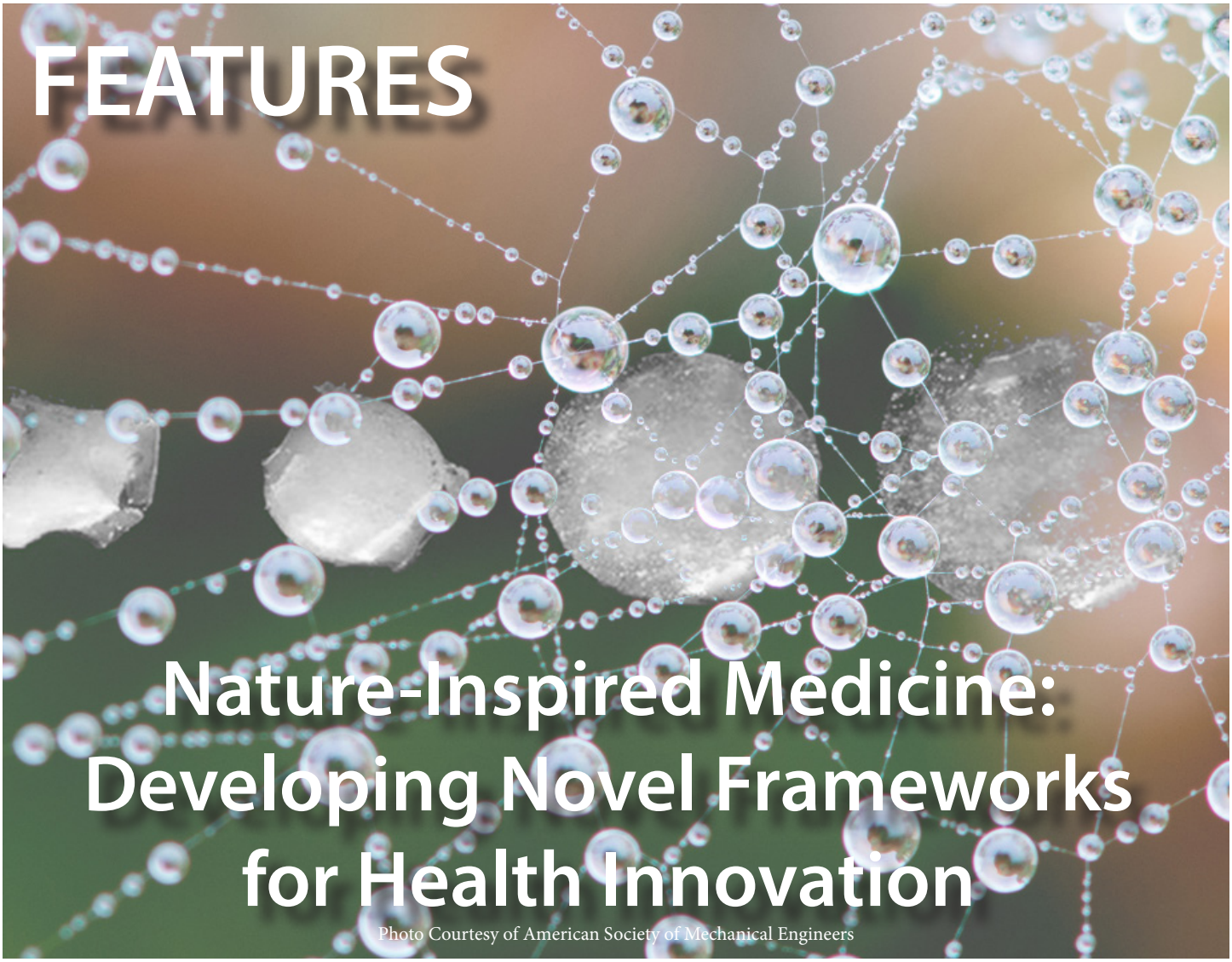
Rivkin, A., Milazzo, M., Venkatesan, A., Frank, E., Vidaurri, M. R., Metzger, P., & Lewicki, C. (2021). Asteroid Resource Utilization: Ethical Concerns and Progress. *Bulletin of the AAS, 53(4).* https://doi.org/10.3847/25c2cfeb.173706d5.

Sanchez, J. A., Reddy, V., Bottke, W. F., Battle, A., Sharkey, B., Kareta, T., Pearson, N., & Cantillo, D. C. (2021). Physical Characterization of Metal-rich Near-Earth Asteroids 6178 (1986 DA) and 2016 ED85. *The Planetary Science Journal, 2(5), 205.* https://doi.org/10.3847/psj/ac235f.

Times, T. M. (2020, April 7). *Russia Compares Trump's Space Mining Order to Colonialism*. The Moscow Times. https://www.themoscowtimes.com/2020/04/07/russia-compares-trumps-space-mining-order-to-colonialism-a69901.

UNOOSA. (1966, December 19). *Outer Space Treaty.* United Nations Office for Outer Space Affairs. https://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/outerspacetreaty.html.

US Congress. (2015). *H.R.2262 - 114th Congress (2015-2016): U.S. Commercial Space Launch Competitiveness Act.* Congress.gov. https://www.congress.gov/bill/114th-congress/house-bill/2262.



Nature-Inspired Medicine: Developing Novel Frameworks for Health Innovation

Photo Courtesy of American Society of Mechanical Engineers

Shivi Srikanth ’26

The Current State of Biomedicine

From prosthetics to gene therapies, biomedical researchers have long sought to develop novel solutions to human health challenges. However, in the United States, an aging population, increased chronic illness, and contractions in research funding are placing an increasing burden on the healthcare system (Mokdad et al., 2018; Moses et al., 2015). In 2024, national healthcare use rose by 8.2 percent, exceeding the GDP growth rate by three percentage points and surpassing $5 trillion in spending (Keehan et al., 2025). With such high demand, research institutions and pharmaceutical companies are struggling to address translational bottlenecks, as well as issues of scale and sustainability.

The largest disconnect in medical research is the “benchtop-to-bedside” gap—a lack of translatability from innovation to clinical deployment—partly due to an emphasis on novelty over applicability (Woolf, 2008). According to the National Institutes of Health, 80–90 percent of studies are abandoned before clinical trials (Woolf, 2008). Furthermore, large sums of funding are funneled towards technologies that never reach patients (Seyhan, 2019). This inefficiency is described by Eroom’s Law, which states that the cost of developing a new therapeutic doubles every nine years, while production efficiency halves (Scannell et al., 2012). At this rate, economists predict that the pharmaceutical industry will spend over $16 billion to develop a single therapeutic by 2043, compared to the $2.6 billion development cost often cited today (Seyhan, 2019). This trend forces the industry to produce only the most profitable drugs, favoring wealthier markets and neglecting critical disease burdens like bacterial pneumonia, which is the greatest infectious cause of death in children worldwide (Impact Global Health, 2025).

The benchtop-to-bedside gap is not the only hurdle that the biomedical industry faces. In the last few decades, antibiotic resistance, shrinking insurance coverage, and disease heterogeneity have intensified (CDC, 2019; Trosman et al., 2015). These shifts in pathology are intensified by climate change, driving demand for environmentally-conscious medicine (Romanello et al., 2024). This strain on healthcare indicates that conventional frameworks for innovation are becoming insufficient amid growing social and biological complexity.

As a response, global health researchers and agencies have turned to nature as a purveyor of healthcare, seeking answers in complex ecosystems and organisms designed over millennia. Earlier this year, the World Health Organization published a report urging widespread adoption of nature-based solutions across the healthcare supply chain as a strategy for disease prevention (WHO, 2025). Nature-based design has the potential to address deficiencies in biotechnology, laying a foundation for scalable, translational solutions and resource-conscious development.

Fundamentals of Nature-Inspiration

For centuries, societies have drawn inspiration from nature to design effective products and systems. From plant fibers for textiles to homes resembling termite mounds for passive cooling, communities across the globe have built close partnerships with the tree canopies and rapid streams around them (AlAli et al., 2023).

"Beyond reorienting the role of humans in the environmental context, nature-inspired design can leverage the intrinsic efficiency of ecosystems."

In recent years, however, humans have experienced an epistemic disconnection from nature, yielding a decline in emotional well-being and a lack of stewardship towards the environment. A 2002 study found that British schoolchildren more accurately recognized Pokémon characters than common wildlife, consistent with evidence suggesting that people in high-income countries spend over 85 percent of their time indoors (Balmford, 2002; Klepeis et al., 2001).

Scholars have posited various reasons for this disconnection from nature, the most relevant being cognitive and philosophical. Cognitive disconnection suggests that a lack of ecological education and time away from nature-centric lifestyles has caused many to forget how to live in harmony with the land (Beery et al., 2023). Philosophical disconnection points to a divergence between humans and nature due to advances in modern technology (Beery et al., 2023). This dispossession of the human condition from its ecological context has created systems that fail to protect nature from climate change and humans from its related health impacts.

To address these failures, nature-inspired design blurs the stark line between natural and built environments and encourages consumers to evaluate their impact upon planetary health (Whitburn et al., 2019). Conversely, exposure to green spaces is associated with decreased mortality and improved cardiovascular health (Twohig-Bennett & Jones, 2018). While not a direct replacement for being outdoors, grass-inspired flooring, curtains that mimic tree cover, and sun lamps reintroduce nature into everyday life.

Beyond reorienting the role of humans in the environmental context, nature-inspired design can leverage the intrinsic efficiency of ecosystems. For example, researchers at Sorbonne University recently developed wind turbine blades based on insect wing geometries, increasing energy conversion efficiency by 35 percent (Cognet et al., 2017). Furthermore, nature-based solutions push scientists to consider a product's entire life cycle, avoiding resource waste and environmental mismanagement. This is commonly seen in green fashion companies that use plant-based dyes for clothing, abandoning petrochemicals to ensure ethical sourcing and ease of disposal (United Nations Environment Programme, 2023).

Biomedicine and Nature: Past, Present, and Future

In medicine, nature-inspired design offers yet another dimension of benefit: biocompatibility. For products that require integration into the body, harnessing chemical compounds and physical mechanisms that have proven successful in the natural world can simplify translation from laboratory to patient (Joyce et al., 2021). Many cultures have implemented natural principles for health. Among Indigenous peoples such as the Koasati, willow bark has been used for pain relief for millennia—now understood to contain salicin, a close precursor to aspirin (Mahdi et al., 2006; Taylor, 1940). Similarly, in South Africa, the Zulu and Xhosa peoples pioneered the use of aloe vera to facilitate rapid healing and soothe burns

(Fisher et al., 2025). Across South Asia, traditions like Ayurveda have leveraged the antiseptic properties of neem oil to shield against infection (National Research Council Panel on Neem, 1992).

These are generally proven approaches, but relying solely on nature-first fundamentals is unrealistic in Western medicine, where technological interventions for health are widespread and lucrative. Federal agencies like the Food & Drug Administration require strict adherence to clinical trial standards and distribution laws and are unlikely to accept cultural knowledge as an evidentiary basis (Food and Drug Administration, 2016). Furthermore, medical innovation in the United States is heavily influenced by a "technological imperative," described by sociologists as persistent pressure to use available technology once it exists (Tymstra, 1989).

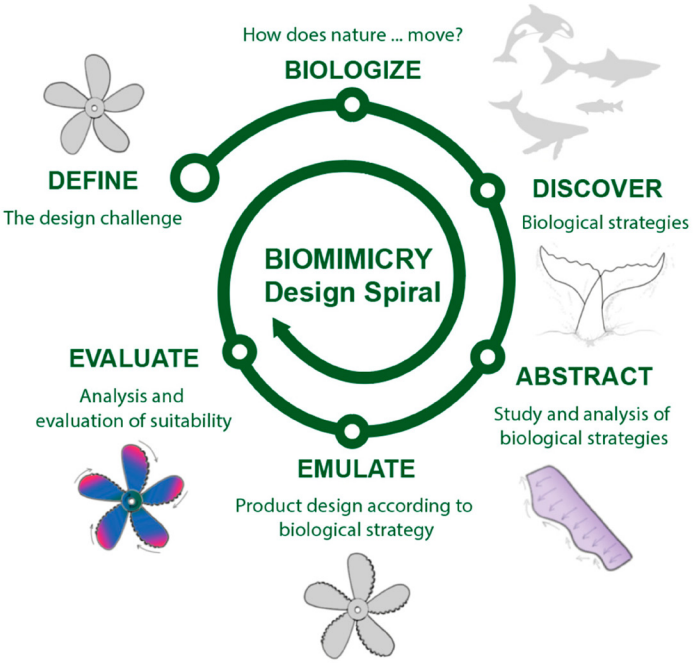


Figure 1. The Biomimicry Spiral. *Reprinted from Aguilar-Planet & Peralta (2024).*

Nature-inspired biomedicine, which retains the integrity of modern medicine and emphasizes an ecological underpinning, offers a compromise.

In some cases, stumbling upon ecological phenomena can lead to novel medical solutions. A classical example is the zebrafish, whose observed regenerative properties launched an entire field of tissue engineering to treat patient-specific disease (Gemberling et al., 2013). However, in the absence of standardized norms to guide sustainable and scalable interventions, it may not be sufficient to rely on chance alone.

To move beyond isolated efforts, frameworks for applying nature-inspired design are gaining traction. The Biomimicry Spiral, as seen in Figure 1, is a tool built by the Biomimicry Institute. Together, the six steps offer a pathway that mirrors standard engineering workflows, while prioritizing ecological considerations at critical junctures (Biomimicry Institute, 2025). Perhaps most dissimilar to standard design models is the "Biologize" function, which pushes researchers to reframe their design question entirely in organismic terms (Biomimicry Institute, 2025).

The gecko is a common model organism in nature-inspired design. Its ability to adhere to vertical surfaces, often regardless of material, has inspired skin-friendly, solvent-free adhesives for wearable devices (Kang et al., 2021). A researcher seeking to design biocompatible bandages might use the "Biologize" function of the Spiral to consider how animals with adhesive properties attach robustly to soft, irregular surfaces. During the "Discover" step, the researcher consults with ecologists to identify natural models that address these challenges and distill the strategies that ensure their success. This knowledge-gathering might uncover the gecko's microscopic underfoot scales as a promising parallel. During the "Abstract" step, the researcher identifies the material and physical properties present in the gecko, like van der Waals force maximization (Kang et al., 2021). The "Emulate" step encourages recapitulation of these properties in novel materials, such as polymer films patterned with microarrays that mimic this underfoot structure and function (Kang et al., 2021). To "Evaluate", the researcher tests the validity of this technology against its original design goals, ensuring successful translation first from organism to product, then from product to patient.

Projects that abide by the Biomimicry Spiral can be compared across similar criteria, where explicit parameters allow different labs to run the same assays and build a shared evidence base instead of producing isolated prototypes. A common framework for sustainable and biocompatible design can drive scalability and an interconnectedness not just between nature and innovation, but between institutions (Damschroder et al., 2009).

Moving Forward: Implementation and Ethics

Nature-inspired design addresses a diverse set of research challenges, making it an attractive framework to adopt. However, implementation poses a substantive hurdle. Short R&D funding cycles hinder long-term, future-facing innovation, and disciplinary silos make

interaction between biomedical researchers and ecologists minimal (Franssen et al., 2018; Manlove et al., 2016).

Despite these structural barriers in the pharmaceutical sector and widespread polarization around climate change, sustainability is becoming an undercurrent in biomedicine (Kennedy & Tyson, 2024). In 2024, nearly 500 U.S. hospitals reported cutting about 185,000 metric tons of emissions, signaling a shift in environmental awareness (Practice GreenHealth, 2025). In addition, the medical biomimetics market is projected to grow by 7.6 percent between 2024 and 2030, suggesting increased receptiveness to the adoption of nature-based strategies (Grandview Research, 2024).

Moving forward, the value of this framework is dependent upon the ethics that govern it. Some scholars argue that nature-based design has the potential to perpetuate systems of extraction and exploitation of Indigenous medical knowledge for Western benefit (Alum, 2024). In 1997, the South African Council for Scientific and Industrial Research patented an appetite-suppressant from a cactus-like succulent (*Hoodia gordonii*), using knowledge long held by indigenous San communities without explicit permission. The Council only negotiated a benefit-sharing agreement, which provided credit and a share of the profits, following years of public pressure (Vermeylen, 2007). In response to cases like these, a 2024 treaty adopted by the World Intellectual Property Organization requires patent applicants to disclose inspiration drawn from indigenous knowledge in an attempt to deter biopiracy and boost accountability (World Intellectual Property Organization, 2024). However, while this treaty signals increasing support for conscious innovation, the United States and United Kingdom—two large patent producers—have yet to sign it.

With this context, any implementation of nature-inspired solutions should not be separated from local impacts. For example, the "Biologize" step of the Biomimicry Spiral must emphasize cultural context alongside function, so that models and materials are codesigned with communities whose knowledge precedes modern medicine. Preventing the extraction of Indigenous knowledge requires that explicit consent, attribution, and benefit-sharing be built into solutions at the moment of conception (Alum, 2024). Nature-inspired design has the potential to rectify inefficiencies within biomedicine and prioritize sustainability ahead of innovation, but only if wielded with legislative guardrails and humility towards the people and ecosystems who have pioneered it.

References

Aguilar-Planet, T., & Peralta, E. (2024). Innovation inspired by nature: Applications of biomimicry in engineering design. *Biomimetics*, *9*(9), 523. https://doi.org/10.3390/biomimetics9090523

AlAli, M., Mattar, Y., Mhd Amer Alzaim, & Salwa Beheiry. (2023). Applications of biomimicry in architecture, construction and civil engineering. *Applications of Biomimicry in Architecture, Construction and Civil Engineering*, *8*(2), 202–202. https://doi.org/10.3390/biomimetics8020202

Alum, E. U. (2024). The role of indigenous knowledge in advancing the therapeutic use of medicinal plants: Challenges and opportunities. *Plant Signaling & Behavior*, *19*(1). https://doi.org/10.1080/15592324.2024.2439255

Balmford, A. (2002). Why conservationists should heed Pokemon. *Science*, *295*(5564), 2367b2367. https://doi.org/10.1126/science.295.5564.2367b

Beery, T., Stahl Olafsson, A., Gentin, S., Maurer, M., Stålhammar, S., Albert, C., Bieling, C., Buijs, A., Fagerholm, N., Garcia-Martin, M., Plieninger, T., & M. Raymond, C. (2023). Disconnection from nature: Expanding our understanding of human–nature relations. *People and Nature*, *5*(2). https://doi.org/10.1002/pan3.10451

Biomimicry Institute. (2025). *The Biomimicry Process*. Biomimicry Toolbox. https://toolbox.biomimicry.org/methods/process/

Centers for Disease Control and Prevention. (2019). *Antibiotic resistance threats report*. https://www.cdc.gov/antimicrobial-resistance/data-research/threats/index.html

Cognet, V., Courrech du Pont, S., Dobrev, I., Massouh, F., & Thiria, B. (2017). Bioinspired turbine blades offer new perspectives for wind energy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *473*(2198), 20160726. https://doi.org/10.1098/rspa.2016.0726

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, *4*(1). https://doi.org/10.1186/1748-5908-4-50

Fisher (née Rahiman), F., Africa, C., Klaasen, J., & Fisher, R. (2025). South African medicinal plants traditionally used for wound treatment: An ethnobotanical systematic review. *Plants*, *14*(5), 818. https://doi.org/10.3390/plants14050818

Food and Drug Administration. (2016, October 6). *Adequate and well-controlled studies*. https://www.ecfr.gov/current/title-21/section-314.126

Franssen, T., Scholten, W., Hessels, L. K., & de Rijcke, S. (2018). The drawbacks of project funding for epistemic innovation: Comparing institutional affordances and constraints of different types of research funding. *Minerva*, *56*(1), 11–33. https://doi.org/10.1007/s11024-017-9338-9

Gemberling, M., Bailey, T. J., Hyde, D. R., & Poss, K. D. (2013). The zebrafish as a model for complex tissue regeneration. *Trends in Genetics*, *29*(11), 611–620. https://doi.org/10.1016/j.tig.2013.07.003

Grandview Research. (2024). *Medical Biomimetics Market (2024 - 2030)*. https://www.grandviewresearch.com/industry-analysis/medical-biomimetics-market-report

Holland, C., Numata, K., Rnjak-Kovacina, J., & Seib, F. P. (2018). The biomedical use of silk: Past, present, future. *Advanced Healthcare Materials*, *8*(1), 1800465. https://doi.org/10.1002/adhm.201800465

Impact Global Health. (2025). *Smart decisions: The G-FINDER 2024 neglected disease R&D report*. https://cdn.impactglobalhealth.org/media/G-FINDER%202024_Full%20report.pdf

Joyce, K., Fabra, G. T., Bozkurt, Y., & Pandit, A. (2021). Bioactive potential of natural

biomaterials: Identification, retention and assessment of biological properties. *Signal Transduction and Targeted Therapy*, *6*(1), 1–28. https://doi.org/10.1038/s41392-021-00512-8

Kang, M., Sun, K., Seong, M., Hwang, I., Jang, H., Park, S., Choi, G., Lee, S.-H., Kim, J., & Jeong, H. E. (2021). Applications of bioinspired reversible dry and wet adhesives: A review. *Frontiers in Mechanical Engineering*, *7*. https://doi.org/10.3389/fmech.2021.668262

Keehan, S. P., Madison, A. J., Poisal, J. A., Cuckler, G. A., Smith, S. D., Sisko, A. M., Fiore, J. A., & Rennie, K. E. (2025). National health expenditure projections, 2024–33: Despite insurance coverage declines, health to grow as share of GDP. *Health Affairs*, *44*(7). https://doi.org/10.1377/hlthaff.2025.00545

Kelly, C. (2021, January 7). *9 bioinspired medical technologies - ASME*. www.asme.org. https://www.asme.org/topics-resources/content/9-bioinspired-medical-technologies

Kennedy, B., & Tyson, A. (2024, December 9). *How americans view climate change and policies to address the issue*. Pew Research Center. https://www.pewresearch.org/science/2024/12/09/how-americans-view-climate-change-and-policies-to-address-the-issue/

Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., Behar, J., Hern, S., & Engelmann, W. (2001). The national human activity pattern survey (NHAPS): A resource for assessing exposure to environmental pollutants. *Journal of Exposure Science & Environmental Epidemiology*, *11*(3), 231–252. https://doi.org/10.1038/sj.jea.7500165

Mahdi, J. G., Mahdi, A. J., Mahdi, A. J., & Bowen, I. D. (2006). The historical analysis of aspirin discovery, its relation to the willow tree and antiproliferative and anticancer potential. *Cell Proliferation*, *39*(2), 147–155. https://doi.org/10.1111/j.1365-2184.2006.00377.x

Manlove, K. R., Walker, J. G., Craft, M. E., Huyvaert, K. P., Joseph, M. B., Miller, R. S., Nol, P., Patyk, K. A., O'Brien, D., Walsh, D. P., & Cross, P. C. (2016). "One health" or three? Publication silos among the one health disciplines. *PLoS Biology*, *14*(4), e1002448. https://doi.org/10.1371/journal.pbio.1002448

Mokdad, A. H., Ballestros, K., Echko, M., Glenn, S., Olsen, H. E., Mullany, E., Lee, A., Khan, A. R., Ahmadi, A., Ferrari, A. J., Kasaeian, A., Werdecker, A., Carter, A., Zipkin, B., Sartorius, B., Serdar, B., Sykes, B. L., Troeger, C., Fitzmaurice, C., & Rehm, C. D. (2018). The state of US health, 1990-2016. *JAMA*, *319*(14), 1444. https://doi.org/10.1001/jama.2018.0158

Moses, H., Matheson, D. H. M., Cairns-Smith, S., George, B. P., Palisch, C., & Dorsey, E. R. (2015). The anatomy of medical research. *JAMA*, *313*(2), 174. https://doi.org/10.1001/jama.2014.15939

National Research Council (US) Panel on Neem. (1992). Medicinals. In *www.ncbi.nlm.nih.gov*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK234637/

Practice GreenHealth. (2025). *The power of partner data: from sector trends to individualized insights*. Practicegreenhealth.org. https://practicegreenhealth.org/about/news/power-partner-data-sector-trends-individualized-insights

Romanello, M., Walawender, M., Hsu, S.-C., Annalyse Moskeland, Yasna Palmeiro-Silva, Scamman, D., Ali, Z., Ameli, N., Denitsa Angelova, Ayeb-Karlsson, S., Basart, S., Beagley, J., Beggs, P. J., Blanco-Villafuerte, L., Cai, W., Callaghan, M., Diarmid Campbell-Lendrum, Chambers, J. D., Chicmana-Zapata, V., & Chu, L. (2024). The 2024 report of the Lancet countdown on health and climate change: Facing record-breaking threats from delayed action. *The Lancet*, *0*(0). https://doi.org/10.1016/S0140-6736(24)01822-1

Scannell, J. W., Blanckley, A., Boldon, H., & Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery*, *11*(3), 191–200. https://doi.org/10.1038/nrd3681

Seyhan, A. A. (2019). Lost in translation: The valley of death across preclinical and clinical divide – identification of problems and overcoming obstacles. *Translational Medicine Communications*, *4*(1). https://doi.org/10.1186/s41231-019-0050-7

Taylor, L. Averill Paz. (1940). Plants used as curatives by certain southeastern tribes. Cambridge, Mass.: Botanical museum of Harvard university.

Trosman, J. R., Weldon, C. B., Kelley, R. K., & Phillips, K. A. (2015). Challenges of coverage policy development for next-generation tumor sequencing panels: Experts and payers weigh in. *Journal of the National Comprehensive Cancer Network*, *13*(3), 311–318. https://doi.org/10.6004/jnccn.2015.0043

Twohig-Bennett, C., & Jones, A. (2018). The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environmental Research*, *166*(1), 628–637. https://doi.org/10.1016/j.envres.2018.06.030

Tymstra, T. (1989). The imperative character of medical technology and the meaning of "anticipated decision regret." *International Journal of Technology Assessment in Health Care*, *5*(2), 207–213. https://doi.org/10.1017/s0266462300006437

United Nations Environment Programme. (2023). *Sustainability and circularity in the textile value chain: A global roadmap*. https://www.unep.org/resources/publication/sustainability-and-circularity-textile-value-chain-global-roadmap

Vermeylen, S. (2007). Contextualizing "fair" and "equitable": The San's reflections on the Hoodia benefit-sharing agreement. *Local Environment*, *12*(4), 423–436. https://doi.org/10.1080/13549830701495252

Whitburn, J., Linklater, W., & Abrahamse, W. (2019). Meta-analysis of human connection to nature and proenvironmental behavior. *Conservation Biology*, *34*(1). https://doi.org/10.1111/cobi.13381

Woolf, S. H. (2008). The meaning of translational research and why it matters. *JAMA*, *299*(2). https://doi.org/10.1001/jama.2007.26

World Health Organization. (2025). *Nature-based solutions and health*. https://www.who.int/europe/publications/i/item/WHO-EURO-2025-12214-51986-79744

World Intellectual Property Organization. (2024). *WIPO Treaty on Intellectual Property, Genetic Resources and Associated Traditional Knowledge* (adopted May 24, 2024). WIPO Lex (Text No. 593055). https://www.wipo.int/wipolex/en/text/593055

Neglect and Excess: The Two Faces of Maternal Health Inequity in South Asia

Photo courtesy of World Bank Photo

Nazifa Ahmed '29

Introduction

South Asia's maternal mortality story is one of both success and stubborn inequity. As the world's most populous region—and home to nearly a third of all global maternal deaths—it offers urgent insight into how inequity can persist even amid progress. Between 2000 and 2023, the region achieved a 71% decline in its maternal mortality rate, from about 405 to 117 maternal deaths per 100,000 live births, one of the steepest reductions globally (World Health Organization, 2025). However, that progress is deeply unequal across socioeconomic status. Women in remote areas, lower-income households, and marginalized communities continue to face preventable death during childbirth due to gaps in care. At the same time, wealthier urban women increasingly experience the opposite problem: the over-medicalization of birth, including the widespread implementation of unnecessary caesarean sections (C-sections). Fragmented health systems in South Asia, driven by unequal public investment and unregulated private expansion, create a dual crisis in maternal care: the neglect of marginalized women and the exploitation of those with access to private medicine.

Underuse of Life-saving Care: Neglect

Since public investments remain focused on urban hospitals while rural facilities lack resources, many women in poorer and remote communities still struggle to obtain even the most basic maternal care. Improvements in survival have largely benefited women who are wealthier, urban, and more socially privileged. A recent study found that women in the lowest income quintile in South Asia are two to three times less likely to receive key maternal services—such as antenatal care, skilled birth attendance, or postpartum checkups—compared to their wealthiest and urban counterparts (Rahman et al., 2024). In Bangladesh, increases in hospital births have occurred mainly among affluent, urban populations, while women in remote districts still rely on home deliveries with minimal support (Nagesh et al., 2024). A similar urban bias emerges across borders: The Pakistan Demographic and Health Survey (PDHS) highlights enduring rural–urban disparities in access to emergency obstetric services (Mumtaz et al., 2017). Even within India, these inequalities

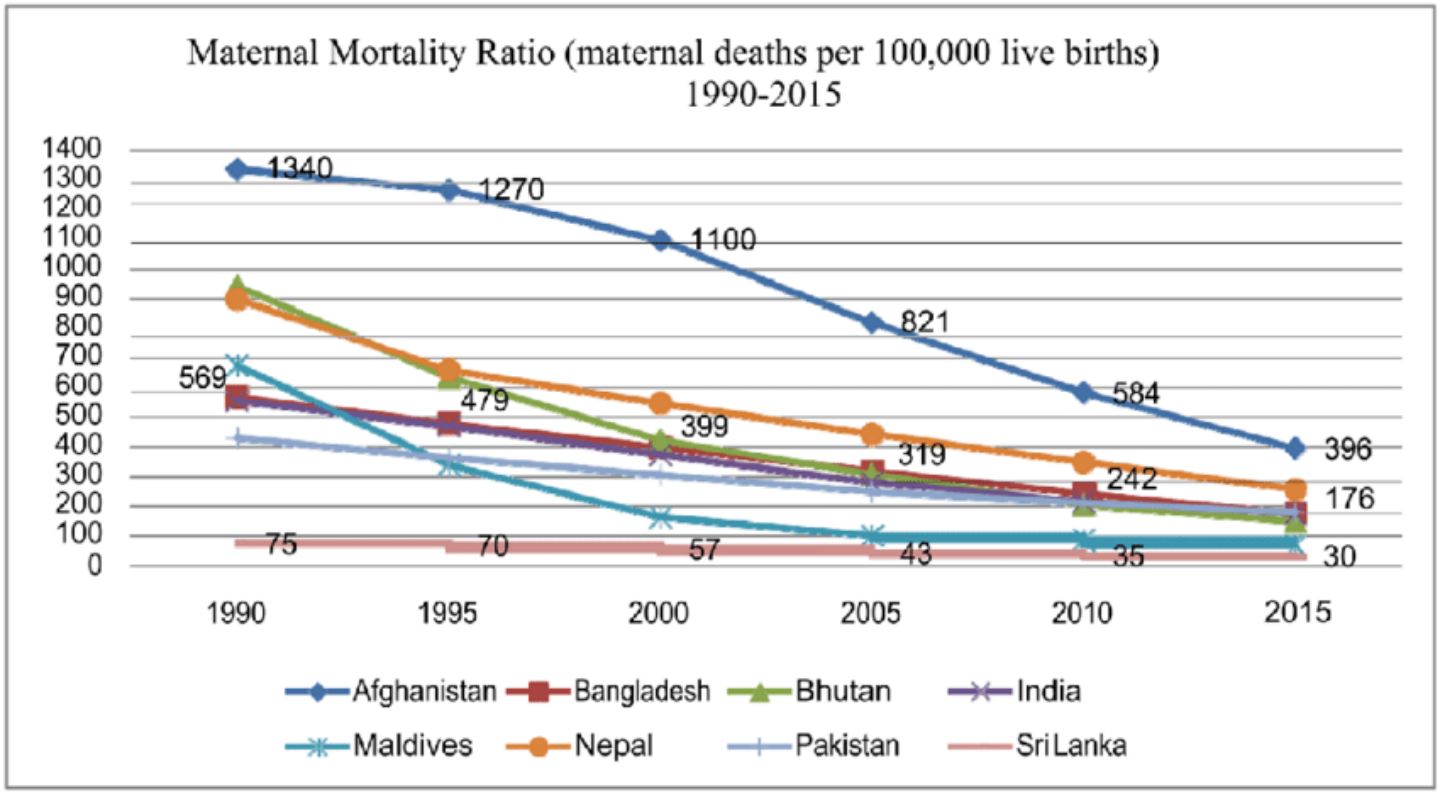


Figure 1. Trends in maternal mortality ratios across South Asia (1990–2015): Despite substantial regional declines, the figure shows wide cross-country disparities driven by differences in public investment, health system capacity, and social inequality. The contrast between low-mortality countries (Sri Lanka, Maldives) and high-mortality ones (Afghanistan, Pakistan) mirrors the structural inequities at the center of South Asia's maternal health landscape (WHO, 2015).

persist at the state level. The Sample Registration System (SRS) Special Bulletin (2020–22) shows that the state of Assam—one of India's poorest states by per-capita income—has a maternal mortality ratio nearly three times higher than Kerala, a southern state known for its robust public health infrastructure and high female literacy. These differences in treatment are not incidental; they reflect broader structural and institutional inequalities (Dutta et al., 2023).

The principal causes of maternal death in South Asia—postpartum hemorrhage, hypertensive disorders such as pre-eclampsia and eclampsia, obstructed labor, sepsis, and indirect sources like anemia, cardiovascular disease—are well-documented (Souza et al., 2024). In principle, these conditions are preventable or treatable with interventions such as uterotonics, safe blood transfusion, magnesium sulfate therapy, timely surgical intervention, antibiotics, and quality antenatal screening. However, they are often unavailable or inaccessible when they are most needed—especially for the women at highest risk.

A helpful framework for understanding this phenomenon is that of the "three delays": delay in seeking care, reaching care, and receiving appropriate care upon arrival. These delays are exacerbated by poverty, inadequate transportation, weak referral networks, gendered decision-making constraints, and limited health system capacity (El-Saharty et al., 2015).

Importantly, underuse is not simply the absence of care. Rather, it also implicates the poor quality of routine services. In some facilities, staff shortages, stockouts of essential supplies (e.g., blood and magnesium sulfate), or lack of surgical capacity turn "institutional delivery" into a false promise of safety, given that these deficiencies result in a reduced capacity to manage emergencies. Reports from parts of Bihar, Sindh, and rural India document this inequity within facilities (Nagesh et al., 2024). Thus, the first face of maternal health inequity in South Asia is structural neglect: when geography, socioeconomic status, caste, and gender govern whether a woman receives timely, life-saving care. As a result of this disparity, it is vital to redesign referral networks, strengthen quality and emergency readiness, and tackle social and logistical barriers.

Overuse of Unnecessary Care: Exploitation

C-section rates have surged across South Asia's urban and private hospitals, often far exceeding medical need. In Bangladesh, for example, C-sections climbed from

just 3.5% of births in 2004 to 32.8% by 2017. Pakistan saw a similar rise from 2.7% to 22.3% over two decades (Rana et al., 2024). These increases are concentrated among wealthy, urban populations: C-section usage in Bangladeshi and Pakistani cities is roughly double that of rural areas (approximately 27% versus 13%; Rana et al., 2024). Private facilities amplify this trend. In Bangladesh, women delivering in private hospitals have nearly five times the odds of receiving a C-section compared to those in public facilities; in Nepal, the odds are roughly twice as high (Rana et al., 2024). India likewise exhibits stark disparities, with nearly half of the deliveries in private hospitals being C-sections versus roughly 14% in public facilities (Dutta et al., 2023). As a reference point, the World Health Organization has suggested an "ideal" C-section rate of around 10%-15% for optimal outcomes. The fact that this threshold has now been surpassed in many South Asian settings suggests that C-sections are increasingly being driven by non-medical factors rather than genuine clinical need.

Financial incentives and convenience have become major drivers of rising C-section rates in private urban care. Underpaid doctors may push elective C-sections to increase income, and scheduled operations are more convenient than unpredictable labor (Rana et al., 2024).

"However, they are often unavailable or inaccessible when they are most needed—especially for the women at highest risk. "

Among educated and affluent women, C-sections are often seen as more modern, safer, and less painful than a vaginal delivery and are sometimes even considered a mark of status and high-quality care (Filipovic, 2014). These misconceptions can lead women to prefer unnecessary surgery (Rana et al., 2024). Family influences and media portrayals —such as television dramas and YouTube birth vlogs that frame C-sections as modern, painless, and convenient— also contribute to a cultural shift toward "planned" births (Veparala et al., 2025).

Though life-saving when indicated, unwarranted surgeries heighten risks of infection, hemorrhage, anesthesia complications, and future pregnancy problems. One analysis by Negrini et al. (2021) noted that unwarranted C-sections can triple the risk of maternal death and lead to a higher incidence of placental disorders in subsequent pregnancies. Babies delivered by C-section without an evident need may suffer respiratory issues and have higher chances of neonatal ICU admission (Negrini et al., 2021).

Economically, C-sections cost far more than vaginal births. In the U.S., they cost about 30% more, with similar disparities in South Asia where most families pay out-of-pocket (Negrini et al., 2021). Such unnecessary procedures deepen inequities: Affluent families pay high prices for interventions of limited benefit, while scarce health funds are diverted from improving essential maternal care. Beyond the financial consequences, over-medicalization also erodes trust. The suspicion that providers may recommend surgeries for profit or convenience rather than medical need can fuel deep mistrust and dissatisfaction (Rana et al., 2024). On a broader scale, communities may become wary of institutional care, weakening the patient–provider relationship and discouraging women from seeking care altogether.

Structural Causes of the Two-Sided Inequity

The coexistence of underuse and overuse of maternal healthcare in South Asia is rooted in deeper structural flaws. As aforementioned, on one side, poor rural women receive too little care; on the other, affluent urban women are overtreated. Most South Asian governments allocate very little public funding to healthcare, resulting in weak rural infrastructure. Public health spending in the region is among the lowest in the world—for example, Bangladesh spends only about 0.4% of its gross domestic product on health, whereas even a small country like the Maldives spends about 8% (Bloch, 2020). Across countries like India, Pakistan, Nepal, and Bangladesh, government funding covers barely half, or less, of total health expenditures, leaving households to pay the rest out-of-pocket (Bloch, 2020). As a result, many rural clinics lack staff, supplies, or emergency obstetric care, and poor women often give birth in the absence of skilled attendants. In contrast, countries like Sri Lanka and the Maldives, which prioritize public health investment, have achieved much lower maternal mortality rates (Bloch, 2020). In the Maldives, where the maternal mortality ratio has declined to around 32 per 100,000 live births, and in Sri Lanka, at about 18 per 100,000 live births, strong public-health investment correlates with lower maternal death rates (World Bank Blog, 2023).

Meanwhile, the rapid but weakly regulated growth of the private sector has transformed urban maternity care. Concentrated in cities and catering to those who can pay, private hospitals often prioritize profit over patient welfare. A recent review by Veparala et al.

(2025) identifies private sector dominance and weak regulatory frameworks as key system-level drivers of rising C-section rates across South Asia. Private hospitals often have financial incentives to offer more procedures and face few penalties for excessive intervention, and insurance schemes may further encourage surgical deliveries by reimbursing them at higher rates. The result is an oversupply of care in privileged settings—wealthy, medically low-risk women subjected to unnecessary procedures driven by commercial rather than clinical motives.

"Regardless, women's autonomy remains constrained: The poor are denied adequate care, while the wealthy are steered towards needless procedures. "

Deep-rooted social hierarchies determine who gets neglected and who gets overtreated. Gender norms in South Asia have historically undervalued women's health, especially for the poor. In rural areas, women from lower castes or disadvantaged ethnic groups are further marginalized—not only because caste discrimination limits their access to quality care, but also because caste and class often overlap, concentrating poverty among these groups and leaving them dependent on under-resourced facilities or home delivery. This Pattern aligns with former Harvard physician-anthropologist Paul Farmer's (2010) concept of structural violence, where social structures (e.g., poverty, inequality, patriarchy) systematically harm certain groups' health. The notion that normal childbirth is for the poor, while the elite opt for surgery has been observed not only in South Asia but also in other countries, capturing how birth and quality of treatment have become a marker of status rather than a medical need (Filipovic, 2014). Regardless, women's autonomy remains constrained: The poor are denied adequate care, while the wealthy are steered towards needless procedures.

Conclusion

Overall, these dynamics expose a dual injustice at the heart of South Asia's maternal health landscape: the neglect of marginalized women and the over-medicalization of those with greater means. Similar inequities appear wherever health systems are segmented. For instance, Brazil's maternal care system shows a familiar split: Low-income women depend on the public sector (with higher rates of natural birth), whereas private hospitals serving the rich have C-section rates of 80% to 90%, far above the recommended 15% (Filipovic, 2014). There, too, C-sections have become "modern and elegant," and vaginal birth is disparaged as something "poor women are supposed to endure" (Filipovic, 2014). By contrast, countries that have strong, well-regulated public health systems manage to avoid such extremes. Many high-income nations ensure more uniform maternity care; for example, in much of Western Europe, robust public provisioning and clinical guidelines keep overall C-section rates moderate and minimize gaps between population groups. Even within South Asia, Sri Lanka stands out as an example of more equitable maternal healthcare. Long-term public investment—including rural midwife networks and free universal maternity care—has led to high coverage of essential services and relatively low maternal mortality (Bloch, 2020). These comparisons underscore that South Asia's two-sided maternal health crisis is not inevitable, but rather the predictable result of policy choices and power structures. Adopting a structural lens—what Farmer (2010) terms a *"pathology of power"*—reveals that the exploitation of some women and the neglect of others are two symptoms of the same underlying problem: a health system stratified by wealth, geography, and social status. Addressing this will require systemic reforms, including greater public investment in underserved areas, regulation of private care to curb profiteering, and efforts to challenge the gender and class biases that have become deeply ingrained into the fabric of the health system.

References

Bloch, C. (2020). *Public spending on health, education and social assistance in South Asian countries*. https://ipcid.org/sites/default/files/pub/en/OP455_Public_spending_on_health_education.pdf

Dutta, R., Nathani, P., Patil, P., Ghoshal, R., Tuli, S., Bakker, J. M., van Duinen, A. J., Roy, N., Boatin, A. A., & Gadgil, A. (2025). State-wise variation and inequalities in caesarean delivery rates in India: Analysis of the National Family Health Survey-5 (2019–2021) data. *The Lancet Regional Health - Southeast Asia*, *32*, 100512. https://doi.org/10.1016/j.lansea.2024.100512

El-saharty, S. (2015). South Asia's Quest for Reduced Maternal Mortality: What the Data *Show. South Asia's Quest for Reduced Maternal Mortality: What the Data Show*. https://blogs.worldbank.org/en/health/south-asia-s-quest-reduced-maternal-mortality-what-data-show? / *South asia's quest for reduced maternal mortality: What the data show*. (n.d.). World Bank Blogs. Retrieved November 2, 2025, from https://blogs.worldbank.org/en/health/south-asia-s-quest-reduced-maternal-mortality-

what-data-show

Farmer, P. (2010a). *Pathologies of power: Health, human rights, and the new war on the poor* (Nachdr.). University of California Press.

Filipovic, J. (2014, April 10). Inside a war on natural birth: C-sections as status symbol and "choice" as a myth. *The Guardian*. https://www.theguardian.com/commentisfree/2014/apr/10/natural-birth-c-section-choice-brazil-forced-pregnancy

Mumtaz, S., Bahk, J., & Khang, Y.-H. (2017). Rising trends and inequalities in cesarean section rates in Pakistan: Evidence from Pakistan Demographic and Health Surveys, 1990-2013. *PLOS ONE*, *12*(10), e0186563. https://doi.org/10.1371/journal.pone.0186563

Nagesh, N., Ip, C. H. L., Li, J., Fan, H. S. L., Chai, H. S., Fan, Y., Wong, J. Y., Fong, D. Y., & Lok, K. Y.-W. (2024). Exploring South Asian women's perspectives and experiences of maternity care services: A qualitative evidence synthesis. *Women and Birth*, *37*(2), 259–277. https://doi.org/10.1016/j.wombi.2023.12.002

Negrini, R., da Silva Ferreira, R. D., & Guimarães, D. Z. (2021a). Value-based care in obstetrics: Comparison between vaginal birth and caesarean section. *BMC Pregnancy and Childbirth*, *21*(1), 333. https://doi.org/10.1186/s12884-021-03798-2

Rahman, M. A., Islam, M. A., Tohan, M. M., Muhibullah, S. M., Rahman, M. S., & Howlader, M. H. (2024). Socioeconomic inequalities in utilizing maternal health care in five South Asian countries: A decomposition analysis. *PloS One*, *19*(2), e0296762. https://doi.org/10.1371/journal.pone.0296762

Rana, M. S., Mazumder, S., Khan, M. T. F., Khan, M. M. H., & Rahman, M. M. (2024). Trends and determinants of caesarean section in south asian countries: Bangladesh, nepal, and pakistan. *PLOS ONE*, *19*(12), e0311082. https://doi.org/10.1371/journal.pone.0311082

Souza, J. P., Day, L. T., Rezende-Gomes, A. C., Zhang, J., Mori, R., Baguiya, A., Jayaratne, K., Osoti, A., Vogel, J. P., Campbell, O., Mugerwa, K. Y., Lumbiganon, P., Tunçalp, Ö., Cresswell, J., Say, L., Moran, A. C., & Oladapo, O. T. (2024). A global analysis of the determinants of maternal health and transitions in maternal mortality. *The Lancet Global Health*, *12*(2), e306–e316. https://doi.org/10.1016/S2214-109X(23)00468-0

Veparala, A., Lall, D., Srinivas, P., Samantaray, K., & Marchal, B. (2025a). Health system drivers of caesarean deliveries in south Asia: A scoping review. *Lancet Regional Health - Southeast Asia*, *40*(100651). https://doi.org/10.1016/j.lansea.2025.100651

World Health Organization. (2025, April 7). *Maternal mortality*. https://www.who.int/news-room/fact-sheets/detail/maternal-mortality?

World Health Organization WHO, United Nations Fund for Children, United Nations Population Fund, World Bank Group and the United Nations Population Division (2015). *Trends in Maternal Mortality: 1990 to 2015.*

"WHO Statement on Caesarean Section Rates." Accessed November 16, 2025. https://www.who.int/publications/i/item/WHO-RHR-15.02.

Hijacking the Powerhouse: Mitochondrial Transfer and Metabolism in Cancer Cell Proliferation

Courtesy of Hoover et al. (2025)

Iphia Zhang ’29

Introduction

Mitochondria are cellular organelles that play a vital role in sustaining life: They influence the fate of stem cells, release signaling molecules, regulate cell death, and modulate energy metabolism through the production of adenosine triphosphate (ATP), which cells metabolize for energy, among other functions (Wang et al., 2024). Due to their utility in maintaining cellular health, mitochondria can drive disease when such processes are disrupted, as can be the case in cancer cells, where the breakdown of cell cycle regulators causes uncontrolled cell division (National Cancer Institute, 2021). As a result, mitochondria may be valuable therapeutic targets to increase cancer cells’ vulnerability to treatment (Mukherjee et al., 2023), given that cell division has a high energy cost, driving the need for increased ATP production in mitochondria. Cancer cells also face low oxygen, low blood supply, and acidic conditions in the tumor microenvironment (TME), the network of immune cells and blood vessels surrounding tumors, forcing cancer cells to reprogram their biology to survive (Fadaka et al., 2017). Uncovering how cancer cells achieve this reprogramming may generate more targeted therapies (Liu et al., 2023).

One way cancer cells compensate for increased energy demands is by taking mitochondria from healthy cells in a process called horizontal mitochondrial transfer (Artusa et al., 2025). Among a population of healthy cells, horizontal mitochondrial transfer can be essential in helping boost metabolism in energy-deficient cells to support their survival, as has been observed in certain neurons and heart muscle cells (Borcherding & Brestoff, 2023). Cancer cells can exploit this process to enhance their own survival. There are several mechanisms by

which cancer cells accomplish horizontal mitochondrial transfer, including through the formation of structures called tunneling nanotubes (TNTs), transport sacs called extracellular vesicles, and cell fusion (Marabitti et al., 2024).

Cancer cells can also promote their own survival using mitochondria through metabolic alterations. Reprogramming of the pathway used for energy generation has been observed in cancer cells alongside changes in other biochemical pathways (Wang et al., 2023). Therefore, whether by physically transferring mitochondria to cancer cells or reprogramming mitochondrial function in cancer cells, mitochondria can effectively be hijacked by cancer cells to sustain their proliferation. These mechanisms demonstrate the adaptability of cancer cells, though further study will be needed to exploit them therapeutically. Targeting mitochondria may help increase the effectiveness of existing drugs or prevent cancer cell survival.

Mitochondrial Transfer to Cancer Cells

Cancer cells must increase their supply of energy in order to meet the demands of rapid cell division and survive in the acidic, low-oxygen TME (Fadaka et al., 2017). The idea that mitochondrial transfer can help recipient cells meet their energy demands has been demonstrated experimentally. Growing cells with nonfunctional mitochondria with cells that had sufficient mitochondria restored respiration in mitochondria-deficient cells (Spees et al., 2006). Additionally, studies have shown that breast cancer cells can become significantly more resistant to chemotherapy upon receiving transferred mitochondria. These data suggest that mitochondria transfer may also help cancer cells become more resistant to treatment and aid the recovery of cancer cells with damaged mitochondria, allowing for their continued survival (Pasquier et al., 2013).

"Mitochondria can effectively be hijacked by cancer cells to sustain their proliferation."

Mitochondrial transfer from healthy cells to cancer cells has been observed in various cell types and cancers. In an experiment where lung cancer cells from mice were grown with immune cells that are normally responsible for eliminating tumors, researchers found that mitochondria from the immune cells moved into the cytoplasm of the cancer cells (Zhang et al., 2023). Cancer cells may also be dependent on mitochondria from nerves. When grown with neurons, cancer cells can acquire mitochondria from neurons, and experimentally cutting off the nerve supply in breast cancer cells can cause slowed tumor growth. Accordingly, in human prostate cancer, cells that are closer to a nerve supply have more mitochondria than those that are without a nerve supply (Hoover et al., 2025). Mitochondria can also be transferred in blood cancers; leukemia cells can transfer mitochondria from cells in the connective tissue and cells that line blood vessels (Moschoi et al., 2016; Pasquier et al., 2013). Mitochondrial transfer may be a significant survival strategy for cancer cells. Thus, preventing the process may weaken their survival and have therapeutic effects (Artusa et al., 2025).

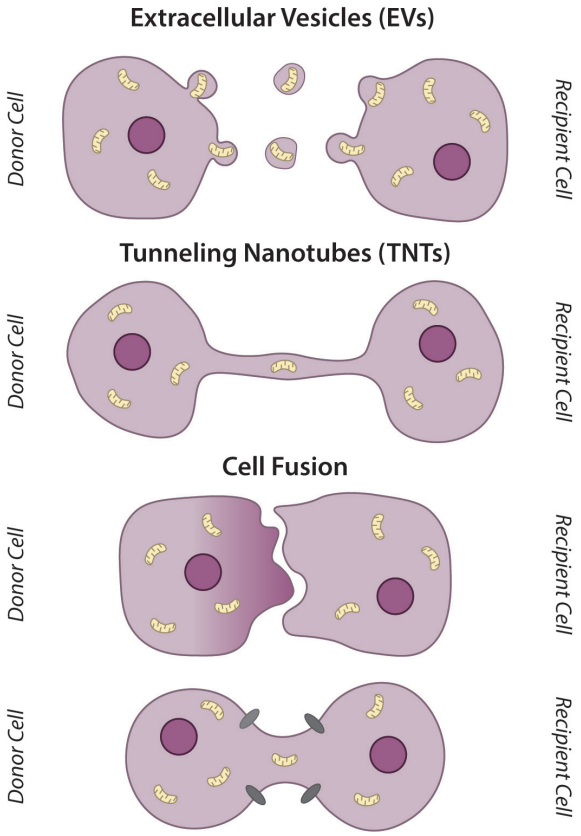


Figure 1. Routes of mitochondrial transfer. (Adapted to include TNTs, EVs, and cell fusion and display the physical elements) (Marabitti et al., 2024b).

Mechanisms of Acquisition

Cancer cells can acquire mitochondria from healthy cells in various ways. For instance, cancer cells are able to make distant, physical connections with nearby

cells by forming a bridge-like protrusion, linking them with healthy cells. Through such protrusions, called tunneling nanotubes, organelles like mitochondria can be transferred (Pasquier et al., 2013). Another way cancer cells can obtain mitochondria is via extracellular vesicles, which package bioactive molecules and deliver them between cells. Through different signaling pathways, cancer cells can traffic mitochondria towards them (Marabitti et al., 2024). Finally, mitochondria are also able to move directly between cancer and healthy cells by cell fusion, where part of the membranes of the cells combine (Spees et al., 2006).

The TME promotes the mechanisms by which cancer cells acquire mitochondria from healthy cells. Stressors like low oxygen cause healthy cells to package their mitochondria, and in the low-oxygen TME, cancer cells may exploit this tendency to acquire mitochondria (Zhang et al., 2021). It has also been seen that ovarian cancer cells that are resistant to chemotherapy form more TNTs in low-oxygen environments like the TME. Thus, low oxygen appears to induce signaling of TNT-forming pathways (Desir et al., 2016). The TME also supports tumor proliferation by suppressing the mitochondrial function of immune cells responsible for attacking tumors, allowing tumor cells to survive and contribute to this immunosuppression by taking mitochondria from the immune cells (Scharping et al., 2016). Because cancer cells rely significantly on TNTs for mitochondrial transfer, devising effective ways to inhibit TNT formation could potentially advance cancer therapies. Several drugs that block structural elements necessary for the process have been identified (Marabitti et al., 2024a).

Alterations in Mitochondrial Metabolism

Inside cancer cells, mitochondria are manipulated in a phenomenon known as the Warburg effect. German physiologist Otto Warburg observed in the 1920s that cancer cells used oxygen-independent glycolysis to generate ATP even in oxygen-sufficient conditions. This shift in metabolism appears to be counterproductive, as the ATP yield from glycolysis is less than that from oxygen-dependent respiration, which takes place in mitochondria (Zhang et al., 2023). Cancer cells tend to favor glycolysis for its faster production of ATP and because it generates intermediate building blocks necessary to support tumor growth. Mitochondria in cancer have been seen to use a different starting material in the energy-generating Krebs cycle, putting intermediates through pathways generating biomolecules needed for cancer proliferation (Wang et al., 2023).

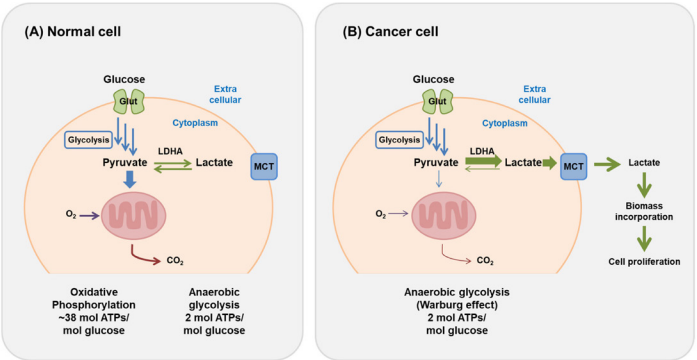


Figure 2. Warburg effect; pyruvate is a product of glycolysis that is normally put through mitochondrial respiration (Kim & Baek, 2021).

This hallmark metabolic reprogramming in some cancers may separate cancer cells from healthy cells in treatment. By inhibiting altered metabolism in reprogrammed cells, more selective targets may be provided. Cancers like acute myeloid leukemia and glioma have mitochondria that harbor mutant enzymes of the Krebs cycle, which can regulate gene expression to promote cancer growth. These cancers are found in TMEs that are resistant to standard anti-cancer drugs, so targeting such mutations may improve disease outcome (Zhang et al., 2023). Cancer cells are also distinguished from healthy cells in their increased levels of antioxidants, which help them balance the increased stress from reactive molecules generated by processes in the mitochondria. In a mouse model of skin cancer, increasing antioxidants promoted the spread of cancer, demonstrating the need for a balance in reactive molecules and antioxidants. Cancer cells could be targeted by blocking pathways that contribute to antioxidant production, subjecting them to the stress placed upon them by increased biochemical activity in mitochondria (Vyas et al., 2016).

Conclusion

Although they hold great therapeutic potential, the exact biochemical mechanisms for mitochondrial manipulations remain unclear. What genes are involved in the formation of TNTs? What are the deeper molecular mechanisms enabling mitochondrial transfer? How does the Warburg effect influence other metabolic processes in cancer cells? Answering these questions may drive the development of new therapies and further the field of cancer biomedicine.

The hijacking of the cellular powerhouse, both by horizontal mitochondrial transfer and by metabolic manipulation, creates a distinct angle from which cancer proliferation can be studied. Understanding how cancer

hijacks mitochondria will allow medicine to combat the numerous side effects brought about by traditional chemotherapy, a consequence of these drugs' inability to distinguish cancer cells from rapidly dividing healthy cells. Identifying areas where cancer cells differ from healthy cells, and what cancer cells rely on for survival, is vital in making therapies more selective and effective, minimizing side effects, and preserving the health of normal cells. The mitochondrial biology of cancer cells is among these hallmarks, and has the potential to advance treatment by reducing cancer's resistance or eradicating its energy supply altogether.

References

Artusa, V., De Luca, L., Clerici, M., & Trabattoni, D. (2025). Connecting the dots: Mitochondrial transfer in immunity, inflammation, and cancer. *Immunology Letters*, *274*, 106992. https://doi.org/10.1016/j.imlet.2025.106992

Borcherding, N., & Brestoff, J. R. (2023). The power and potential of mitochondria transfer. *Nature*, *623*(7986), 283–291. https://doi.org/10.1038/s41586-023-06537-z

Desir, S., Dickson, E. L., Vogel, R. I., Thayanithy, V., Wong, P., Teoh, D., Geller, M. A., Steer, C. J., Subramanian, S., & Lou, E. (2016). Tunneling nanotube formation is stimulated by hypoxia in ovarian cancer cells. *Oncotarget*, *7*(28). https://doi.org/10.18632/oncotarget.9504

Fadaka, A., Ajiboye, B., Ojo, O., Adewale, O., Olayide, I., & Emuowhochere, R. (2017). Biology of glucose metabolization in cancer cells. *Journal of Oncological Sciences*, *3*(2), 45–51. https://doi.org/10.1016/j.jons.2017.06.002

Grelet, S., & Ayala, G. (2025). Cancer cells (red, with their nuclei stained blue) grow thin tubes to siphon mitochondria out of nerve cells (green). In *Nature*. https://media.nature.com/w1248/magazine-assets/d41586-025-01941-z/d41586-025-01941-z_51140130.jpg?as=webp

Hoover, G., Gilbert, S., Curley, O., Obellianne, C., Lin, M. T., Hixson, W., Pierce, T. W., Andrews, J. F., Alexeyev, M. F., Ding, Y., Bu, P., Behbod, F., Medina, D., Chang, J. T., Ayala, G., & Grelet, S. (2025). Nerve-to-cancer transfer of mitochondria during cancer metastasis. *Nature*, *644*. https://doi.org/10.1038/s41586-025-09176-8

Kim, S. H., & Baek, K. H. (2021). Regulation of Cancer Metabolism by Deubiquitinating Enzymes: The Warburg Effect. *International Journal of Molecular Sciences*, *22*(12), 6173.

Liu, J., Zhang, J., Gao, Y., Jiang, Y., Guan, Z., Xie, Y., Hu, J., & Chen, J. (2023). Barrier permeation and improved nanomedicine delivery in tumor microenvironments. *Cancer Letters*, *562*, 216166–216166. https://doi.org/10.1016/j.canlet.2023.216166

Marabitti, V., Vulpis, E., Nazio, F., & Campello, S. (2024a). Mitochondrial Transfer as a Strategy for Enhancing Cancer Cell Fitness: Current Insights and Future Directions. *Pharmacological Research*, *208*, 107382–107382. https://doi.org/10.1016/j.phrs.2024.107382

Marabitti, V., Vulpis, E., Nazio, F., & Campello, S. (2024b). Routes of Mitochondrial Transfer. In *Pharmacological Research* (Vol. 208, pp. 107382–107382). Elsevier BV. https://doi.org/10.1016/j.phrs.2024.107382

Moschoi, R., Imbert, V., Nebout, M., Chiche, J., Mary, D., Prebet, T., Saland, E., Castellano, R., Pouyet, L., Collette, Y., Vey, N., Chabannon, C., Recher, C., Sarry, J.-E. ., Alcor, D., Peyron, J.-F. ., & Griessinger, E. (2016). Protective mitochondrial transfer from bone marrow stromal cells to acute myeloid leukemic cells during chemotherapy. *Blood*, 128(2), 253–264. https://doi.org/10.1182/blood-2015-07-655860

Mukherjee, S., Gurjit Kaur Bhatti, Chhabra, R., P. Hemachandra Reddy, & Jasvinder Singh Bhatti. (2023). Targeting mitochondria as a potential therapeutic strategy against chemoresistance in cancer. *Biomedicine & Pharmacotherapy*, *160*, 114398–114398. https://doi.org/10.1016/j.biopha.2023.114398

National Cancer Institute. (2021, October 11). *What Is Cancer?* National Cancer Institute; National Institutes of Health. https://www.cancer.gov/about-cancer/understanding/what-is-cancer

National Cancer Institute. (2025). *Chemotherapy to Treat Cancer*. National Cancer Institute; National Cancer Institute. https://www.cancer.gov/about-cancer/treatment/types/chemotherapy

Oncodaily. (2024). Warburg effect. In *Oncodaily*.

Pasquier, J., Guerrouahen, B. S., Al Thawadi, H., Ghiabi, P., Maleki, M., Abu-Kaoud, N., Jacob, A., Mirshahi, M., Galas, L., Rafii, S., Le Foll, F., & Rafii, A. (2013). Preferential transfer of mitochondria from endothelial to cancer cells through tunneling nanotubes modulates chemoresistance. *Journal of Translational Medicine*, *11*(1), 94. https://doi.org/10.1186/1479-5876-11-94

Scharping, N. E., Menk, A. V., Moreci, R. S., Whetstone, R. D., Dadey, R. E., Watkins, S. C., Ferris, R. L., & Delgoffe, G. M. (2016). The Tumor Microenvironment Represses T Cell Mitochondrial Biogenesis to Drive Intratumoral T Cell Metabolic Insufficiency and Dysfunction. *Immunity*, *45*(2), 374–388. https://doi.org/10.1016/j.immuni.2016.07.009

Spees, J. L., Olson, S. D., Whitney, M. J., & Prockop, D. J. (2006). Mitochondrial transfer between cells can rescue aerobic respiration. *Proceedings of the National Academy of Sciences*, *103*(5), 1283–1288. https://doi.org/10.1073/pnas.0510511103

Vyas, S., Zaganjor, E., & Haigis, M. C. (2016). Mitochondria and Cancer. *Cell*, *166*(3), 555–566. https://doi.org/10.1016/j.cell.2016.07.002

Wang, Q., Yuan, Y., Liu, J., Li, C., & Jiang, X. (2024). The role of mitochondria in aging, cell death, and tumor immunity. *Frontiers in Immunology*, *15*. https://doi.org/10.3389/fimmu.2024.1520072

Wang, S.-F., Tseng, L., & Hsin Chen Lee. (2023). Role of mitochondrial alterations in human cancer progression and cancer immunity. *Journal of Biomedical Science*, *30*(1). https://doi.org/10.1186/s12929-023-00956-w

Zhang, H., Yu, X., Ye, J., Li, H., Hu, J., Tan, Y., Fang, Y., Akbay, E., Yu, F., Weng, C., Sankaran, V. G., Bachoo, R. M., Maher, E., Minna, J., Zhang, A., & Li, B. (2023). Systematic investigation of mitochondrial transfer between cancer cells and T cells at single-cell resolution. *Cancer Cell*, *41*(10), 1788-1802.e10. https://doi.org/10.1016/j.ccell.2023.09.003

Zhang, L., Wei, Y., Yuan, S., & Sun, L. (2023). Targeting mitochondrial metabolic reprogramming as a potential approach for cancer therapy. International Journal of *Molecular Sciences*, *24*(5), 4954. https://doi.org/10.3390/ijms24054954

Zhang, Y., Tan, J., Miao, Y., & Zhang, Q. (2021). The effect of extracellular vesicles on the regulation of mitochondria under hypoxia. *Cell Death & Disease*, *12*(4). https://doi.org/10.1038/s41419-021-03640-9

FEATURES

Medicine in Captivity: How Internment Shaped Japanese American Health For Generations

Photo courtesy of Densho Encyclopedia

Clara Shin ’29

Introduction

Eighty-three years later, the internment of Japanese Americans during World War II (WWII; 1939–1945) remains one of the most egregious violations of civil liberties in the history of the United States. In the wake of the 1941 attack on Pearl Harbor, the U.S. government failed to distinguish between imperial Japan as a foreign power and Japanese Americans (Office of the Historian, n.d.). As a result, widespread xenophobia and wartime hysteria fueled suspicion toward anyone of Japanese ancestry.

Following the signing of Executive Order 9066 in 1942, approximately 120,000 Japanese Americans were forcibly removed from their homes and relocated to internment camps (National Archives, 1942). The conditions in the camps were harsh: remote locations with extreme temperatures, inadequate food supplies, and poor sanitation. Edward J. Ennis, then a Member of the Board of Directors of the American Civil Liberties Union, condemned the ongoing mass incarceration as “the greatest deprivation of liberty since slavery” (Ennis, 1984).

One of the most telling indicators of the quality of life in these camps was the state of healthcare. While healthcare today is recognized as a fundamental human right—formally articulated in the World Health Organization’s 1946 Constitution—U.S. policy had not incorporated such a concept during WWII. Thus, within the camps, patients receiving medical care were also forced to contend with chronic shortages of medicine, equipment, and trained physicians. Although the U.S. government assigned some medical personnel to the camps, these staff were few in number and often inadequately supplied. Interned medical professionals—many of whom had been practicing doctors, nurses, or pharmacists before their incarceration—worked under severe constraints, receiving little to no pay and lacking

the necessary resources to provide adequate care (Nagata & Takeshita, 1998). Although the U.S. government made some efforts to offer medical services, such as sending meager supplies, these attempts were largely insufficient, and often led to preventable illness, suffering, and death (Nakayama & Jensen, 2011).

Overall, these failures of healthcare in the camps were a reflection of racialized policies that determined access to basic human needs. Moreover, the effects of this medical neglect stretch far beyond the closure of the camps; its consequences have lingered for generations, affecting the physical and mental health of Japanese American families in contemporary society.

Healthcare in the Internment Camps

The forced removal of Japanese Americans dismantled entire healthcare networks that had served these communities, particularly in urban areas. Many Japanese American doctors, nurses, and pharmacists lost their practices, and community hospitals were shuttered overnight (Nakayama & Jensen, 2011).

Internees were initially sent to makeshift assembly centers, where living conditions were deplorable (Nagata & Takeshita, 1998). Families were crammed into overcrowded barracks, including converted horse stalls that reeked of manure, while rudimentary sanitation facilities led to frequent outbreaks of disease (Jensen, 2008). Overcrowding and poor ventilation created the perfect conditions for the rapid spread of tuberculosis and influenza. These crude infirmaries lacked essentials such as intravenous needles, tubing, and fluids. Even with emergency supplies, the Japanese medical staff struggled to treat the overwhelming number of patients (Fiset, 2024). Additionally, although medical records are sparse, oral histories frequently describe cases in which treatable physical illnesses such as tuberculosis or the flu were misdiagnosed, and "melancholia" or medical complaints were dismissed as somatic symptoms of psychological distress (Dusselier, 2008). As Dr. Yoshiko "Fred" Fujikawa, an interned Japanese American physician, later recalled from the Santa Anita assembly center in California, "people lined up to go to the latrine and […] some of them fainted in the line" (Nakayama & Jensen, 2011). His recollection exemplifies the exhaustion of the first phase of internment.

"The deprivation of basic necessities and systemic neglect that the internees faced in the camps created a lasting imprint on the health of survivors."

Even upon transfer to the more permanent War Relocation Authority camps, the healthcare remained inadequate. These camps often lacked basic resources like ambulances, surgical equipment, and adequate prenatal care. Hospitals, if constructed at all, took months to open. Japanese American physicians and nurses received as little as $19 per month and, despite their professional training, were routinely overruled by White supervisors, who often abused their authority (Fiset, 2024).

The racial hierarchy within healthcare systems at the camps added to the stress for both patients and providers. White medical directors often discounted the expertise of more experienced Japanese American doctors. For instance, Dr. Reece Pedicord, the White medical director at the Tule Lake internment camp in Northern California, grew suspicious that Japanese American physicians were performing an excessive number of surgeries (Nakayama & Jensen, 2011). Dr. Shigeru Hara, an interned Japanese physician, remembered that "he [Pedicord] used to say, 'Any case that has to be operated [on], you've got to see me first.' Some of those were acute appendicitis. By the time he got to see them, he had made it pretty late and made it much worse. So that's why a lot of people were against him" (Nakayama & Jensen, 2011). This dynamic reveals how deeply anti-Japanese prejudice was woven into the structure of American medicine: the hierarchy at the camps mirrored the broader racialized exclusion of minorities within American healthcare, which ultimately exacerbated patient suffering.

Long-Term Health Consequences of Internment

The health consequences of Japanese American internment went far beyond the years spent in the camps. The deprivation of basic necessities and systemic neglect that the internees faced in the camps created a lasting imprint on the health of survivors.

In the camps, the widespread malnutrition and overcrowding created an environment in which both infectious and chronic illnesses flourished. Chronic conditions such as diabetes, ulcers, and cardiovascular disease were often poorly managed in the camps, due to a lack of medications, specialized care, and consistent monitoring, leaving some with irreversible complications that continued into adulthood (Jensen, 2005). In some

cases, previously mild illnesses worsened irreversibly due to neglect or stress-induced flare-ups. For example, tuberculosis, which thrived in the overcrowded and poorly ventilated barracks, could cause permanent lung damage (Jensen, 2005). The spread of coccidioidomycosis (also known as Valley fever) was also exacerbated by the dusty, arid locations of many of the camps (Fiset, 2024). The lack of prenatal care for pregnant internees added further risks: miscarriages, preterm births, and birth complications were more frequent, and mothers often lacked access to trained medical staff during labor.

These physical consequences were compounded by severe psychological trauma. Internment involved not only the loss of homes, but also communities, autonomy, and dignity—all of which generated widespread symptoms consistent with those of post-traumatic stress disorder (Cai & Lee, 2022). Depression and anxiety were especially common among older internees who had lived stable lives before the war.

Intergenerational Effects of Internment

While survivors endured significant trauma during the internment, its impact also affected the mental and emotional health of later generations. Research indicates that third-generation Japanese Americans, or Sansei, whose parents experienced incarceration during their own childhoods, exhibited higher levels of distress and lower subjective well-being later in life compared to non-interned Japanese Americans (Nagata et al., 1999). These effects echo the findings from intergenerational trauma studies across multiple populations, such as the descendants of Holocaust survivors, Cambodian genocide survivors, and Indigenous boarding school attendees (Isobel et al., 2021).

Moreover, the internment experience often disrupted traditional family structure and caregiving roles. Children of internment survivors who witnessed their parents lose control over daily life were found to be at higher risk of developing a fractured sense of identity and intergenerational detachment (Spiel et al., 2023). Broader studies suggest that such environments can lead to insecure attachment styles, which correlate with long-term relational and psychological challenges, such as anxiety, depression, or emotional dysregulation.

The legacy of this internment has also played a crucial role in how trauma was passed down. Many second-generation (Nisei) parents remained silent about their experiences in the camps, often out of shame, a desire to assimilate, or a fear of re-experiencing the trauma. This silence created what scholars call "emotional gaps" between generations, in which children felt the presence of unspoken trauma without being able to name or process it (Isobel et al., 2021). Children often interpreted their parents' abnormal behavior as personal or emotional withdrawal, which compounded their own struggles with identity, anxiety, and belonging.

Research on intergenerational communication in Asian American families supports this dynamic (Hesse & Main, 2000). When parents fail to discuss past trauma, children are left to make sense of it on their own, sometimes internalizing it in ways that harm their mental health, such as self-blame or emotional suppression (Isobel et al., 2021). This process is described by Isobel et al. (2021) as a major mechanism for the transmission of trauma's cross-generational effects.

"The same patterns of systemic neglect and medical mistrust seen in the internment camps persist today in marginalized communities."

In this way, the effects of internment were transmitted through generations. The internalized cultural expectation among Japanese Americans to suppress grief and 'move on' after the war likely intensifies the psychological strain carried by their children (Cai & Lee, 2022). Without space for communal mourning or storytelling, trauma became embedded in their memories and family dynamics. Recognizing this intergenerational trauma is essential not only for an accurate historical understanding but also for informing contemporary approaches to mental health care.

Conclusion

The internment of Japanese Americans during WWII was a clear violation of civil liberties, with the consequences extending across generations. The incarceration camps created an environment of deprivation: overcrowded living quarters, insufficient food, inadequate medical care, and the psychological burden of forced displacement and racial scapegoating. These conditions inflicted immediate harm, and their consequences persisted long after the camps closed.

Attempts to understand the history of Japanese American internment are necessary, not only for Japanese American communities, but for how we understand the

lingering health effects of social injustice wherever it may pervade. The same patterns of systemic neglect and medical mistrust seen in the internment camps persist today in marginalized communities. Black Americans still face unequal treatment in hospitals, Indigenous communities suffer from chronically underfunded health systems, and immigrants in ICE detention experience the same deprivation once normalized in internment camps. The echoes of internment are not just metaphorical; they reveal a recurring national failure to protect the health and dignity of those deemed "other." Recognizing these parallels reminds us that history's injustices will continue to shape who receives healthcare and who is unjustly denied it.

References

Cai, J., & Lee, R. M. (2022). Intergenerational Communication about Historical Trauma in Asian American Families. *Adversity and resilience science*, *3*(3), 233–245. https://doi.org/10.1007/s42844-022-00064-y

Dusselier, J. E. (Jane E. (with Densho). (2008). *Artifacts of loss: Crafting survival in Japanese American concentration camps*. New Brunswick, N.J. : Rutgers University Press. http://archive.org/details/artifactsoflossc00unse

Ennis EJ. Statement in support of The Civil Liberties Act of 1983. Presented to: Subcommittee on Administrative Law and Governmental Relations, Committee on the Judiciary, United States House of Representatives; 1984 Sep 12.

Fiset, Louis. Medical care in camp. (2024, August 29). *Densho Encyclopedia*. Retrieved 17:29, November 2, 2025 from https://encyclopedia.densho.org/Medical%20care%20in%20camp.

Isobel, S., McCloughen, A., Goodyear, M., & Foster, K. (2021). Intergenerational Trauma and Its Relationship to Mental Health Care: A Qualitative Inquiry. *Community mental health journal*, *57*(4), 631–643. https://doi.org/10.1007/s10597-020-00698-1

Jensen, G. M. (n.d.). *Dysentery, dust, and determination: Health care in the World War II Japanese American detention camps*. Discover Nikkei. http://discovernikkei.org/en/journal/2008/6/21/enduring-communities/

Jensen, G. M. (2005). The experience of injustice : health consequences of the Japanese American internment. Thesis (Ph. D.)--University of Colorado, 1997.

Office of the Historian, United States Department of State. (n.d.). *Japanese-American relations at the turn of the century, 1900–1922*. In *Milestones in the History of U.S. Foreign Relations – 1899-1913*. Retrieved from https://history.state.gov/milestones/1899-1913/japanese-relations

Nagata, D. K., & Takeshita, Y. J. (1998). Coping and resilience across generations: Japanese Americans and the World War II internment. *Psychoanalytic review*, *85*(4), 587–613.

Nagata, D. K., Trierweiler, S. J., & Talbot, R. (1999). Long-term effects of internment during early childhood on third-generation Japanese Americans. *The American journal of orthopsychiatry*, *69*(1), 19–29. https://doi.org/10.1037/h0080378

Nakayama, D. K., & Jensen, G. M. (2011). Professionalism behind barbed wire: health care in World War II Japanese-American concentration camps. *Journal of the National Medical Association*, *103*(4), 358–363. https://doi.org/10.1016/s0027-9684(15)30317-5

National Archives and Records Administration. (n.d.). *Executive order 9066: Resulting in Japanese-American incarceration (1942)*. National Archives and Records Administration. https://www.archives.gov/milestone-documents/executive-order-9066

New arrivals vaccinated upon arrival at Manzanar, California, April 2, 1942. (2020, November 18). *Densho Encyclopedia*. Retrieved 21:47, November 23, 2025 from https://encyclopedia.densho.org/sources/en-denshopd-i226-00034-1/.

Spiel, S., Szymanski, K., & Bornstein, R. (2023). Intergenerational Trauma, Dependency, and Detachment. *The Journal of nervous and mental disease*, *211*(9), 679–685. https://doi.org/10.1097/NMD.0000000000001682