Beyond Implicit Bias

Thomas D. Albright, William A. Darity Jr., Diana Dunn, Rayid Ghani, Deena Hayes-Greene, Tanya Katerí Hernández & Sheryl Heron

In their introduction to this edition of *Dædalus*, Goodwin Liu and Camara Phyllis Jones write that "it is unlikely that implicit bias can be effectively addressed by cognitive interventions alone, without broader institutional, legal, and structural reforms." They note that the genesis for the volume was a March 2021 workshop on the science of implicit bias convened by the Committee on Science, Technology, and Law of the National Academies of Sciences, Engineering, and Medicine.[1] That workshop provided an opportunity to demonstrate that implicit bias is a common form of cognitive processing that develops in response to social, cultural, and institutional conditions. As demonstrated by the workshop and the essays in this volume, an understanding of implicit bias in a neurological, mechanistic, and phenomenological manner strengthens our ability to develop policies to diffuse and mitigate the problems that arise from implicit bias.

At the end of the 2021 event, members of the interdisciplinary workshop planning committee gave their perspectives on the important messages that they would take away from the workshop. For the conclusion of this volume of *Dædalus*, we members of the planning committee were asked to expand on what we said three years ago. This is our response.

Thomas D. Albright

Broadly considered, implicit bias is a cognitive response to uncertainty, in which other pieces of information are unconsciously recruited to fill in the blanks of experience. This inferential process is probabilistic and sometimes yields catastrophic outcomes. This is particularly true in a human social context, in which uncertainty is pervasive and other pieces of information include tribal allegiances and social structures that yield disparate treatment as a function of race. Unconscious incorporation of this information allows it to be manifested as implicit racial bias.

The essays in this issue of *Dædalus* catalog the incidence of implicit biases and their devastating effects on individual opportunities and social cohesion. They also explore the societal forces and mechanisms responsible for the development and perseveration of biases. This evidence-based understanding sets the stage for the most important question: how do we stop this from happening?

There are three information-processing strategies that hold promise: 1) reduce uncertainty, 2) change the priors, and 3) compensate. As seen from the essays in this volume, there has been significant growth of science that tests the effectiveness of these strategies.

Reduce uncertainty. Unconscious biases flourish where there is paucity of information. In the case of implicit racial bias, this comes from long-standing forms of cultural and geographic segregation. Evidence suggests that uncertainty can be reduced by engineering meaningful interactions between people from different racial groups, such that information about the "unknown other" is acquired broadly over time and different contexts.

Change the priors. Implicit bias is a form of statistical inference based on observed events and associations. Explicit racism in American society provides a model from which generations of children have acquired a distorted sense of the character of people of different races. As long as that model exists, our priors contain incorrect information, yielding unconscious bias. Hope lies in the fact that acquiring new associations predictably alters unconscious inference, manifested as changes in perception, decision, and action.

Compensate. Efforts to reduce uncertainty and change flawed priors are long-term solutions. Along the way, one valuable strategy is to recognize the biases we hold and overpower them. Because this compensation requires a rational conscious consideration of the potential for error under normal conditions of unconscious bias, the simplest and perhaps most immediately effective strategy is to pause and think before a decision to act. Implicit bias training commonly focuses on this moment, in which qualified decision-making can prevent the harmful biases we have acquired, and hope to suppress, from having real impact on the world in front of us.

William A. ("Sandy") Darity Jr.

The related concepts of unconscious bias and implicit bias have potential value in analyzing personal interactions fraught with prejudice. The two concepts enable individual bigoted acts predicated upon stereotypical beliefs to be viewed as devoid of intent or malice. Both concepts can improve our understanding of interpersonal racism.

In contrast, unconscious bias and implicit bias are far less useful in understanding structural racism, those social practices and policies that produce and

sustain racial inequality. Those practices and policies have been constructed and maintained in both conscious and explicit fashion by their designers.

For example, the incorporation of decentralized authority of the administration of the GI Bill enacted after World War II was an intentional measure to benefit white veterans at the expense of Black veterans. In the late nineteenth century, the failure of the federal government to fulfill its promise of forty-acre land grants to the formerly enslaved as restitution for their years of bondage, while mobilizing the Homestead Act of 1862 to provide 1.5 million white families with 160-acre land grants in the Western territories, was deliberate and purposeful. In the 1950s through the 1970s, the grossly disproportionate placement of freeways under the federal highway system in the heart of predominantly Black neighborhoods and business districts was calculated and willful.

Of course, there have been policies adopted for purposes other than preserving racial hierarchy that have had an inordinate adverse effect on Black community well-being. However, those effects could have been anticipated and mitigated had a careful racial impact audit been performed in advance of their introduction. The failure to conduct such an audit has been the product of a conscious decision by policymakers.

Diana Dunn

In volume one of *Undoing Racism: A Philosophy of International Social Change*, psychologists Ronald Chisholm and Michael Washington define race as "a specious classification of human beings created by Europeans (whites) which assigns human worth and social status using 'white' as the model of humanity and the height of human achievement for the purpose of establishing and maintaining privilege and power."[2]

When white people struggle, we change structures, but when Black people struggle, we give them programs that function within current structures. Only by empowering those most impacted by racism can we create movements that lead to meaningful change.

Rayid Ghani

As we discuss implicit bias in humans, we also need to consider, understand, and deal with that implicit bias propagating in society through the design and use of AI systems. As AI systems augment various existing human processes in society through widespread use by governments and corporations, developing approaches to better understand, detect, and deal with them is a critical need for society.

AI systems have the potential to help improve outcomes and result in a better and more equitable society across a broad range of areas including social welfare, health, education, and criminal justice. At the same time, any AI (or otherwise developed) system that affects people's lives must be explicitly built to increase equity and focus on tackling the implicit (and sometimes explicit) biases underlying the design choices made in the development of that system. It is important to recognize that AI can have a positive social impact, but we need to make sure that we put guidelines, resources, and trainings in place to maximize the chances of the positive impact while protecting people who have historically been negatively impacted by implicit bias in society, and will likely continue to be affected negatively by the new AI systems. This requires government agencies, businesses, nonprofits, and community groups to come together and collaborate around this goal, and for policymakers to act and provide guidelines and/or regulations for both the public- and private-sector organizations using AI-assisted decision-making processes to ensure that these systems are built in a transparent and accountable manner and result in fair and equitable outcomes for society.

When designed appropriately and deliberately, AI can be a useful tool to assist us in both detecting implicit biases when they're present in a human process, as well as in designing new collaborative systems that help reduce the impact of these biases.

Deena Hayes-Greene

Cultural and racial biases are often thought to be an indication of racial prejudice or bigotry – and they can be. However, they are as likely to be the result of the associations our brains make, many of which we are unaware. These associations (for example, shark-danger, fire-hot) are made to save our lives. They become racial when our focus is not on fire or sharks but on people. People whom we might not have met but who are associated with danger or negative labels ascribed by a society, a nation, established in more racially exclusive times. These associations have been nourished for centuries by a culture designed to advantage white people. And the structures such a culture builds and sustains successfully separate and harm poor and working-class people of all races. Drilled into our heads are long-established, time-honored associations about who is valuable, worthy, and deserving. Race then remains the ever-ready tool to prompt well-rooted but implicit associations. Far from the prejudices of individuals, the constant repetition of a hierarchy of human worth, when commonly held, directs the very construction of today's world and creates disparity.

Without interruption, such associations remake themselves with each generation because racism is a toxin in our nation's groundwater. In other words, if you

come upon a lake and see one fish belly up, dead, examining the fish and seeing what caused its death would make sense. If, however, you came upon the lake and half the fish were belly up, what might it be time to do? It is, of course, time to examine the lake. Fish represent the individuals for whom we care, and lakes represent institutions whose toxins could well have caused the need for care. This is possible because lakes are not stand-alone bodies of water. They are fed by the groundwater. The groundwater is the unseen water that makes up 95 percent of the fresh water on our planet. When infected by racism, the groundwater carries it to many lakes, causing many problems. Short of that understanding, our decisions, interventions, and even our visions are misguided, as they fail to see the depth of the problem.

Were we to remove the toxic racial structures underpinning our society, no racial associations would be made. Racial stratification would no longer exist, and racial disparity would be a relic of the past. Racism would no longer inform policy, practice, law, or cultural norms, be our associations explicit or implicit. When we understand not only the brain processes enabling racial associations but also why the associations are there to be made, we can face the past and end its legacy.

Tanya Katerí Hernández

When legal scholars and lawyers consider the literature on implicit bias, they do so for pragmatic reasons. Their principal interest is the desire to enhance the antidiscrimination law project of identifying and addressing discrimination.[3] Indeed, it was the enactment of the Civil Rights Act of 1964's ban on employment discrimination that inspired early iterations of social science–informed workplace trainings.[4]

Contemporary diversity trainings have largely turned to focusing on concerns with implicit bias.[5] However, the trainings for the most part have emphasized the relevance of implicit bias for the individual, and not its implications for structural racism.[6] Yet it is structural racism that antidiscrimination law is geared to address in the context of crafting institutional remedies for findings of discrimination.

It is thus quite heartening that the National Academies' implicit bias workshop not only explained implicit bias, but also linked it to systemic problems. Importantly, two-thirds of human resource specialists report that individual-focused trainings have no effect on the careers of people of color or diversity within the ranks of management, and little effect on levels of implicit bias.[7] Concrete interventions focused on systemic and structural issues make the difference between good and bad implicit bias training.[8]

When training is framed as pertaining to systemic problems and then coupled with complementary measures that engage decision-makers in seeking structur-

al interventions for those systemic problems, workplace diversity is markedly increased as a matter of hiring, retention, and promotion.[9] An example that captured national attention provides a useful illustration. On April 18, 2018, Starbucks employees called the Philadelphia police emergency line to request aid. The cause? Two Black men sitting at a table without placing an order as they waited for the third member of their party to arrive for a meeting. The police arrested the men for "trespassing" and escorted them out of Starbucks in handcuffs. No other White patrons sitting at tables received the same treatment. After cell phone footage of the incident caused a public uproar, Starbucks issued a public apology and closed more than eight thousand U.S. stores for an afternoon of racial bias training for one hundred and seventy-five thousand employees. Notably, the training was accompanied by a structural policy change to disrupt the operation of implicit bias. The new policy states that "any customer is welcome to use Starbucks spaces, including our restrooms, cafes and patios, regardless of whether they make a purchase."[10]

Including concerns about systemic racism in the implicit bias training at Starbucks helped create company support for the structural change with the greatest efficacy for containing the harm of implicit bias. As such, this particular Starbucks effort can serve as a model for how consumers of implicit bias training should encourage program facilitators to speak to systemic and structural issues.[11]

Sheryl Heron

In my professional career, I have been both an emergency physician, where I worked to "stop the bleeding" for patients whom I see for traumatic events, and a public health practitioner, with a focus on injury prevention. While we have been trying to address implicit bias issues at the individual level, the problem is in the prevention of racism at the systemic level. As an immigrant from Jamaica and a Black woman, I call our attention to consider the impact of the caste system. We simply cannot stop at implicit bias or racism. We need to consider the role of caste in this country. Pulitzer Prize–winning author Isabel Wilkerson notes that racism is an insufficient term for the systemic oppression of Black people in America and we must consider the caste system that is a part of this country.[12] In the end, we must ensure we show compassion and empathy in the way we treat one another. We need to move beyond implicit bias, and we must be focused and intentional about how we change the narrative.

ABOUT THE AUTHORS

Thomas D. Albright, a Fellow of the American Academy since 2003, is Professor and Director of the Vision Center Laboratory and Conrad T. Prebys Chair in Vision Research at the Salk Institute for Biological Studies. He has written numerous articles, which have appeared in such journals as *Neuron*, *Journal of Neurophysiology*, and *Proceedings of the National Academy of Sciences of the United States of America*.

William A. ("Sandy") Darity Jr. is the Samuel DuBois Cook Professor of Public Policy, African and African American Studies, and Economics and Director of the Samuel DuBois Cook Center on Social Equity at Duke University. He is the author of *From Here to Equality: Reparations for Black Americans in the Twenty-First Century* (with A. Kirsten Mullen, 2022) and *Persistent Disparity: Race and Economic Inequality in the United States Since 1945* (with Samuel L. Meyers, 1998).

Diana Dunn is the Core Trainer and Organizer at People's Institute for Survival and Beyond.

Rayid Ghani is a Distinguished Career Professor in the Machine Learning Department and the Heinz College of Information Systems and Public Policy at Carnegie Mellon University. He has published in such journals as *Journal of Intelligent Information Systems*, *The British Medical Journal*, and *American Journal of Public Health*.

Deena Hayes-Greene is the Cofounder and Managing Director of Racial Equity Institute, LLC.

Tanya Katerí Hernández is the Archibald R. Murray Professor of Law at Fordham University School of Law, and its Associate Director of the Center on Race, Law, and Justice. She is the author of many articles and the books *Racial Innocence: Unmasking Latino Anti-Black Bias and the Struggle for Equality* (2022), *Multiracials and Civil Rights: Mixed-Race Stories of Discrimination* (2018), and *Racial Subordination in Latin America: The Role of the State, Customary Law and the New Civil Rights Response* (2012).

Sheryl Heron is Professor of Emergency Medicine, Vice-Chair of Faculty Equity, Engagement, and Empowerment for Emergency Medicine, and Chief Diversity and Inclusion Officer, as well as Associate Dean for Community Engagement, Equity, and Inclusion at Emory School of Medicine. She is the editor of *Diversity and Inclusion in Quality Patient Care*, first edition (with Anna Walker Jones, Lisa Moreno-Walton, and Marcus L. Martin, 2016) and second edition (Marcus L. Martin, Lisa Moreno-Walton, and Michelle Strickland, 2019).

ENDNOTES

[1] Goodwin Liu and Camara Phyllis Jones, "Introduction: Implicit Bias in the Context of Structural Racism," *Dædalus* 153 (1) (Winter 2024): 12, https://www.amacad.org/publication/introduction-implicit-bias-context-structural-racism.

[2] Ronald Chisholm and Michael Washington, *Undoing Racism: A Philosophy of International Social Change*, Volume 1 (New York: People's Institute Press, 1997), 30–31.

[3] Charles R. Lawrence III, "The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism," *Stanford Law Review* 39 (1987): 317–388.

4 Lily Zheng, *Deconstructed DEI: Your No-Nonsense Guide to Doing the Work and Doing It Right* (Oakland: Berrett-Koehler Publishers, 2023), 149.

5 Jennifer Y. Kim, "I'm Biased and So Are You. What Should Organizations Do? A Review of Organizational Implicit-Bias Training Programs," *Consulting Psychology Journal* 74 (2022): 19–39.

6 Jesse Singal, *The Quick Fix: Why Fad Psychology Can't Cure Our Social Ills* (New York: Farrar, Straus and Giroux, 2021), 193.

7 Frank Dobbin and Alexandra Kalev, "Why Doesn't Diversity Training Work? The Challenge for Industry and Academia," *Anthropology Now* 10 (2018): 48–55.

8 Alexander Kalev, Frank Dobbin, and Erin Kelly, "Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies," *American Sociological Review* 71 (4) (2006): 589–617.

9 Tanya Katerí Hernández, "Can CRT Save DEI: Workplace Diversity, Equity & Inclusion in the Shadow of Anti-Affirmative Action," *UCLA Law Review Discourse* 71 (forthcoming 2024).

10 Starbucks (@Starbucks), "We want our stores to be the third place, a warm and welcoming environment where customers can gather and connect. Any customer is welcome to use Starbucks spaces, including our restrooms, cafes and patios, regardless of whether they make a purchase. https://sbux.co/2IVVAJ8," Twitter, May 29, 2018, https://twitter.com/Starbucks/status/1001584492249731072.

11 It is important to note that one can appreciate Starbucks's DEI efforts, while at the same time acknowledging the critique of Starbucks's consumers boycotting in opposition to the company's dealings with its employee union and its stance with regards to the crisis in Gaza. Omar Mohammed, "Are McDonald's, Starbucks Boycotts Working?" *Newsweek*, November 17, 2023, https://www.newsweek.com/mcdonalds-starbucks-boycotts-israel-hamas-war-1844933.

12 Isabel Wilkerson, *Caste: The Origins of Our Discontents* (New York: Random House, 2020).