Is This How Discrimination Ends?

Trainings and workshops geared toward eliminating people's hidden prejudices are all the rage—but many don't work. Now the psychologist who made the case for 'implicit bias' wants to cure it.

Story by Jessica Nordell   May 7, 2017 | Science

ON A CLOUDY day in February, Will Cox pointed to a pair of news photos that prompted a room of University of Wisconsin, Madison, graduate students to shift in their seats. In one image, a young African American man clutches a carton of soda under his arm. Dark water swirls around his torso; his yellow shirt is soaked. In the other, a white couple is in water up to their elbows. The woman is tattooed and frowning, gripping a bag of bread.

Cox read aloud the captions that were published alongside these images of a post-Katrina New Orleans. For the black man: "A young man walks through chest-deep water after looting a grocery store." For the white couple: "Two residents wade through chest-deep water after finding bread and soda."

Looting. Finding. A murmur spread through the rows of students watching.

Cox, a social psychologist in the university's Prejudice and Intergroup Relations Lab, turned to his co-presenter, a compact, 50-something woman standing next to him. As she strode down the rows of students, her voice was ardent, her movements deliberate. She could have been under a spotlight on the stage at a tech summit, not at the head of a narrow classroom in the university's education building.

"There are a lot of people who are very sincere in their renunciation of prejudice," she said. "Yet they are vulnerable to habits of mind. Intentions aren't good enough."

The woman, Patricia Devine, is a psychology professor and director of the Prejudice Lab. Thirty years ago, as a graduate student, she conducted a series of experiments that laid out the psychological case for implicit racial bias—the idea, broadly, is that it's possible to act in prejudicial ways while sincerely rejecting prejudiced ideas. She demonstrated that even if people don't believe racist stereotypes are true, those stereotypes, once absorbed, can influence people's behavior without their awareness or intent.

Now, decades after unraveling this phenomenon, Devine wants to find a way to end it. She's not alone. Since the mid-1990s, researchers have been trying to wipe out implicit bias. Over the last several years, "unconscious-bias trainings" have seized Silicon Valley; they are now de rigueur at organizations across the tech world.

But whether these efforts have had any meaningful effect is still largely undetermined.

Until, perhaps, now. I traveled to southern Wisconsin, because Devine and a small group of scientists have developed an approach to bias that actually seems to be working—a two-hour, semi-interactive presentation they've been testing and refining for years. They've created versions focused on race and versions focused on gender. They've tried it with students and faculty. Next, they'll test it with police.

Their goal is to make people act less biased. So far, it's working.

ON JULY 17, 2014, a 43-year-old former horticulturist on Staten Island named Eric Garner was approached by police officers who suspected him of selling untaxed cigarettes. One of them put him in a chokehold, a maneuver the New York City Police Department prohibits. Garner died an hour later—a homicide, according to the medical examiner. Since Garner's death, and then Michael Brown's, and Tamir Rice's, and many, many others', voices condemning discrimination in policing have grown to a thunder. While police in many cases maintain that they used appropriate measures to protect lives and their own personal safety, the concept of implicit bias suggests that in these crucial moments, the officers saw these people not as individuals—a gentle father, an unarmed teenager, a 12-year-old child—but as members of a group they had learned to associate with fear.

In Silicon Valley, a similar narrative of pervasive bias has unfolded over the last several years. In 2012, venture capitalist Ellen Pao filed a gender-discrimination lawsuit against her employer, the venture-capital firm Kleiner

Perkins Caufield & Byers, maintaining, for instance, that she was penalized for the same behaviors her male colleagues were praised for. Her experience wasn't unique: A 2016 survey of hundreds of women in technology, titled Elephant in the Valley, revealed that the vast majority experienced both subtle and overt bias in their careers.

In addition to urgent conversations about race and criminal justice, and employment and gender, discussions about implicit bias have spread to Hollywood, the sciences, and the presidential election. What's more, though solutions are hard to come by, there's plenty of hard data to validate a very real problem.

In fact, studies demonstrate bias across nearly every field and for nearly every group of people. If you're Latino, you'll get less pain medication than a white patient. If you're an elderly woman, you'll receive fewer life-saving interventions than an elderly man. If you are a man being evaluated for a job as a lab manager, you will be given more mentorship, judged as more capable, and offered a higher starting salary than if you were a woman. If you are an obese child, your teacher is more likely to assume you're less intelligent than if you were slim. If you are a black student, you are more likely to be punished than a white student behaving the same way.

There are thousands of these studies. And they show that at this moment in time, if person A is white and person B is black, if person X is a woman and person Y is a man, they will be treated differently in American society for no other reason than that their identities have a cultural meaning. And that meaning clings to each person like a film that cannot be peeled away.

Findings like these can feel radioactive. Ben Barres, a Stanford neurobiologist I spoke with, once wondered aloud whether it was wise to even share with women entering science what they are up against; as James Baldwin wrote, it takes a rare person "to risk madness and heartbreak in an attempt to achieve the impossible." For people struggling to grapple with bias, these realities can elicit feelings of rage and sadness, grief and guilt. Last summer, a man from North Carolina called in to C-SPAN because he wanted to know, quite simply, how he could become less racially biased. "What can I do to change?" he asked. "You know, to be a better American?" The video has been watched online more than 8 million times.

What if the photographers really did see one person looting and not the other?

At the same time, there are many people who reject the concept of implicit bias outright. Some misinterpret it as a charge of plain, old-fashioned bigotry; others just don't perceive its existence, or they believe its role in determining outcomes is overstated. Mike Pence, for instance, bristled during the 2016 vice-presidential debate: "Enough of this seeking every opportunity to demean law enforcement broadly by making the accusation of implicit bias whenever tragedy occurs." And two days after the first presidential debate, in which Hillary Clinton proclaimed the need to address implicit bias, Donald Trump asserted that she was "essentially suggesting that everyone, including our police, are basically racist and prejudiced."

Still other people, particularly those who have been the victims of police violence, also reject implicit bias—on the grounds that there's nothing implicit about it at all.

One challenge to bridging these perspectives is that real life rarely provides lab-perfect conditions in which proof of implicit bias can be established. Take Cox's Hurricane Katrina photos. After they were published, people began to ask: What if the photographers really *did* see one person looting and not the other? When the photographers were asked what they'd seen, the photographer of the "looting" photo said that he did see that person loot. The other photographer said that he did not see how his subjects acquired their groceries. There was a plausible explanation other than bias. In debates and jury trials, doubts like this scatter like seeds.

But there may be an even more fundamental reason for this gulf between people's perspectives on the subject of bias. This has to do with the fact that a person's very circumstances and position in the world influence what they do and don't perceive. As Evelyn Carter, a social psychologist at the University of California, Los Angeles, told me, people in the majority and the minority often see two different realities.

While people in the majority may only see intentional acts of discrimination, people in the minority may register both those acts and unintended ones. White people, for instance, might only hear a racist remark, while people of color might register subtler actions, like someone scooting away slightly on a bus—behaviors the majority may not even be aware they're doing.

Bias is woven through culture like a silver cord woven through cloth. In some lights, it's brightly visible. In others, it's hard to distinguish. And your position relative to that glinting thread determines whether you see it at all.

ONE ATTEMPT TO get a handle on all this and pin down implicit bias has come by way of a tool known as the Implicit Association Test. There are many techniques to measure implicit associations, but the IAT, a reaction-time test that gauges how strongly certain concepts are linked in a person's mind, has become the most well-known and widely used.

In an IAT designed to assess anti-gay bias, for instance, you are presented with a list of words (like "smiling" or "rotten" or "homosexual") and asked to sort each into a combined category: "gay or bad" or "straight or good." Then, you see another list and are asked to sort each word again, this time with the combinations flipped: "gay or good" or "straight or bad." If your responses are faster when "gay" is paired with "bad" than with "good," this is supposed to demonstrate that the gay/bad association in your mind is stronger. In a review of more than 2.5 million of these tests, most people showed a preference for straight people over gay, white people over black, and young people over old.

Ideally, the IAT would provide not only a way to quantify bias, but a clear target in the quest to end it. If the implicit associations the IAT measures are the cause of biased behavior, then untethering these negative associations could eliminate it.

But an increasingly vocal group of critics now questions all these assumptions. One issue, according to detractors, is that people's IAT scores are not stable. In scientific parlance, the IAT has low "test-retest reliability"; the same person might end up with different scores at different times. (If a bathroom scale says you weigh 210 pounds today and 160 tomorrow, you might feel skeptical about the scale.) More importantly, according to meta-analyses (a synthesis of all available studies on a topic), there's a weak relationship between a person's IAT score and their actual behavior. In other words, people may be acting in biased ways, but it's not clear this is due to the mental construct measured by the IAT.

The term implicit bias "has become so broad that it almost has no meaning."

While this debate could threaten decades of IAT-based work on implicit bias, a third point of view has also quietly emerged. Yes, the IAT has flaws, this perspective holds. But the meta-analyses criticizing it also have flaws. Furthermore, the IAT is only one of many tools for measuring implicit associations, and all these different tools tend to turn up the same results—the same preferences for certain social groups over others. There *is* something truly consistent and real there, these results suggest.

Perhaps, this third view holds, what's really going on is that implicit bias is more complex than these tools can fully handle. Implicit associations may not be the stable entities scientists and others have been imagining them to be. In fact, studies show that the specific associations that arise depend on a person's context and state of mind. People in one experiment, for instance, automatically associated rich foods with positive things when they were prompted to focus on taste, and negative things when they were prompted to focus on health. In this case, the test-retest reliability criticism would be irrelevant; something that's fluid and changeable *shouldn't* be consistent. The fact that there's any correlation at all between IAT scores and behavior would be remarkable.

All of which is to say that while bias in the world is plainly evident, the exact sequence of mental events that cause it is still a roiling question. Devine, for her part, told me that she is no longer comfortable even calling this phenomenon "implicit bias." Instead, she prefers "unintentional bias." The term implicit bias, she said, "has become so broad that it almost has no meaning."

AN HOUR INTO the workshop in Wisconsin, Devine rolled up one sleeve of her blue paisley blouse and walked over to an African American student sitting in the front row. "People think, 'If I don't want to treat people based on race, then I'm going to be colorblind, or gender-blind, or age-blind,'" she said. "It's not a very effective strategy. First, it's impossible." She held her pale forearm next to the student's. "There's a difference," she said. Students exchanged glances. "Who's the man?" she continued, looking at Cox. Then she raised her eyebrows and gestured to herself. "Who's the old person?" It was a goofy-Dad joke, but the students chuckled.

(How this borderline-aggressive approach affects students put on the spot is another question. Later, the student Devine had approached with her bare arm said, "I was a little surprised, but I kind of appreciated it.")

If pointing out skin color in a workshop on overcoming bias seems strange, that may be part of the point. Trying to ignore these differences, Devine says, makes discrimination worse. Humans see age and gender and skin color: That's vision. Humans have associations about these categories: That's culture. And humans use these associations to make judgments: That, Devine believes, is habit—something you can engage in without knowing it, the way a person might nibble fingernails down to the bloody quick before realizing they are even doing so.

The entire workshop is crafted as a way to break this habit. To do so, Devine said, you have to be aware of it, motivated to change, and have a strategy for replacing it. Over the course of the two-hour presentation, Cox and Devine hit all these notes: They walked through the science of how people can act biased without realizing it. They barreled through mountains of evidence and detailed explanations of how bias spreads. Halfway through the workshop, they gave students a chance to discuss how these ideas relate to their own personal lives, and everyone had a story. A chemist recounted being steered to a sales internship, despite seven years of chemistry training, because she had "such a nice personality." An African American student described the eerie sensation he experienced in Italy—before realizing it was the feeling of *not* having shopkeepers follow and watch him.

At the end, Devine and Cox offered ideas for substitute habits. Observe your own stereotypes and replace them, Cox said. Look for situational reasons for a person's behavior, rather than stereotypes about that person's group. Seek out people who belong to groups unlike your own. Devine paced among the desks, making eye contact with each student. "I submit to you," she said, her voice steady with conviction, "that prejudice is a habit that can be broken."

It may sound a bit whiz-bang, but the data show that this seemingly simple intervention works. For example, after Devine and a colleague presented a version of the workshop focused on gender bias to STEM faculty at the University of Wisconsin, departmental hiring patterns began to change. In the two years following the intervention, in departments that had received the training, the proportion of women faculty hired rose from 32 percent to 47 percent—an increase of almost 50 percent—while in departments that hadn't received the training, the proportion remained flat. And in an independent survey conducted months after the workshop, faculty members in participating departments—both women and men—reported feeling more comfortable bringing up family issues, and even feeling that their research was more valued.

In another study, the researchers gave hundreds of undergraduate students a version of the intervention focused on racial bias. Weeks afterwards, students who had participated noticed bias more in others than did students who hadn't participated, and they were more likely to label the bias they perceived as wrong. Notably, the impact seemed to last: Two years later, students who took part in a public forum on race were more likely to speak out against bias if they had participated in the training.

In treating bias as a habit, the Madison approach is unique. Many psychology experiments that try to change implicit bias treat it as something like blood pressure—a condition that can be adjusted, not a behavior to be overcome. The Madison approach aims to make unconscious patterns conscious and intentional. "The problem is big. It's going to require a variety of different strategies," Devine says. "But if people can address it within themselves, then I think it's a start. If those individuals become part of institutions, they may carry messages forward."

The STEM work suggests this approach at least can have an impact on discrimination against women. What the team has not yet determined is whether the race-focused interventions have an impact on the experiences of people of color. That question is a current priority. "If we're just making white people feel better," Devine says, "who cares?"

Another potential strength of the Madison approach is that it's both rigorously experimental *and* tested in the real world. When the Princeton psychologist Betsy Levy Paluck reviewed hundreds of interventions designed to reduce prejudice, she found that only 11 percent of all experimental efforts were actually tested outside of a laboratory. By contrast, few of the trainings that have become popular in Silicon Valley and elsewhere are scientifically evaluated. This is very concerning to researchers, because the trainings could be having literally any effect on people. In the words of an uncommonly candid acupuncturist I once visited, "After this treatment, you might get better, you might get worse, or you might just stay the same."

Making things worse is a serious risk: A 2006 review of more than 700 companies that had implemented diversity initiatives showed that after diversity training, the likelihood of black men and women advancing in their

organizations actually *decreased*. Proving that efforts like these work as intended, Paluck wrote, "should be considered an ethical imperative, on the level of rigorous testing of medical interventions."

TWO DAYS AFTER the workshop, I sat down in a soaring, glass-walled building on campus with Patrick Forscher, a postdoc in Devine's lab and a co-author on many of the studies evaluating this workshop. I wanted to know why this approach was working.

Forscher is shy; his voice was so low it was almost sub-auditory. Their success, Forscher explained, may have something to do with the creation of the self. In the 1970s, a social psychologist named Milton Rokeach posited that the self is made of many layers and that some layers are more central to one's self-concept than others. Values are highly central; beliefs a little less so. Something like associations would likely be less central still. "If you're asked to describe who you are," said Forscher, "you're more likely to say, 'I'm someone who values equality' than 'I'm someone who implicitly associates white people with good.'"

This hierarchy matters, because the more central a layer is to self-concept, the more resistant it is to change. It's hard, for instance, to alter whether or not a person values the environment. But if you do manage to shift one of these central layers, Forscher explained, the effect is far-reaching. "If you think of therapy, the goal often is to change processes central to how people view themselves," he said. "When it works, it can create very large changes."

The Madison workshop, for its part, zeroes in on people's beliefs about bias—the belief that they might be discriminating, the belief that discrimination is a problem, the belief that they can overcome their own habit of prejudice. As a result of the workshop, these beliefs grow stronger. And beliefs might be just the sweet spot that needs to be targeted. Call beliefs "the Goldilocks layer." They're a central enough part of each person to unleash a torrent of other changes, but removed enough from entrenched core values that they can, with the right kind of pressure, be shifted.

"I don't think it's dark; it's real. You can't pretend it doesn't exist."

Watching Devine, I was struck by how charismatic and funny she is when presenting. It's intentional. If people feel attacked, she told me, they shut down. Avoiding blame is key. The resulting message is carefully balanced: Bias is normal, but it's not acceptable. You must change, but you're not a bad person. Watching Cox and Devine is like watching people play the classic game Operation, tweaking specific beliefs without nicking the wrong reaction.

Still, this approach has shortcomings. A problem as complex as prejudice can't possibly be solved by a single psychological fix. Joelle Emerson, the CEO of Paradigm, a Silicon Valley consultancy that is in the process of evaluating the effectiveness of its own trainings and interventions, believes that long-term change must come through overhauling workplace systems and processes, not relying on individuals.

"Even the most well-designed training is not going to solve things by itself," she said. "You have to reinforce ideas within broader organizations." Social scientists such as the psychologist Glenn Adams have begun to call for a shift "from the task of changing individual hearts and minds to changing the sociocultural worlds in which those hearts and minds are immersed."

There is an elephant in the workshop, too. On the day I attended, almost all the students in the audience were women or people of color, some seeking a way to combat bias directed at them. One student with shoulder-length blond hair confided in me that she cared a lot about these topics, but had hoped the workshop would address what to do about experiencing bias. The absence of white men in the group was conspicuous. Cox said the crowd is usually more mixed, but the audience makeup may point to a fundamental limitation of this kind of work: Its success depends on people already caring enough about these issues to show up in the first place. Not everyone will seek out, in the middle of a weekday, a fairly technical presentation about changing their own habits of mind.

That said, the workshop may come to them anyway. Forscher recently conducted what's known as a "network analysis" of the workshop's effect—a look at how its effects spread throughout a community. What he found, in the STEM gender-bias intervention, was that after the workshop, the people who reported doing the most for gender equity weren't those who'd attended the training, but those who worked alongside them. It's an unusual finding, and it's not clear exactly what this means. But it's possible that as people who attended the workshop changed, they began influencing the people around them.

When Devine first developed the idea that hidden stereotypes can influence people's behavior without their awareness, a colleague observed that her work revealed "the dark side of the mind." Devine disagreed. "I actually think that it reveals the mind," she said. "I don't think it's dark; it's real. You can't pretend it doesn't exist." Neural connections aren't moral. What people do with them is. And as Devine sees it, that doing starts with awareness.

And if there's one thing the Madison workshops do truly shift, it is people's concern that discrimination is a widespread and serious problem. As people become more concerned, the data show, their awareness of bias in the world grows, too. In the days after attending, I noticed my own spontaneous reactions to other people to an almost overwhelming degree. The day I left Madison, in the lobby of my hotel, I saw two people standing near the front desk. They were wearing worn, rumpled clothes, with ragged holes in the knees. As I glanced at them, a story about them flickered in my mind. They weren't guests staying here, I thought; they must be friends of the clerk, visiting him on his break.

It was a tiny story, a minor assumption, but that's how bias starts: as a flicker—unseen, unchecked—that taps at behaviors, reactions, and thoughts. And this tiny story flitted through my mind for seconds before I caught it. *My God,* I thought, *is this how I've been living?*

Afterwards, I kept watching for that flutter, like a person with a net in hand waiting for a dragonfly. And I caught it, many times. Maybe this is the beginning of how my own prejudice ends. Watching for it. Catching it and holding it up to the light. Releasing it. Watching for it again.

JESSICA NORDELL *is a writer based in Minnesota.*