

# Exposing unconscious bias

How much of the prejudice that shapes our worlds is unconscious, and can we truly measure it, asks **Pragya Agarwal**

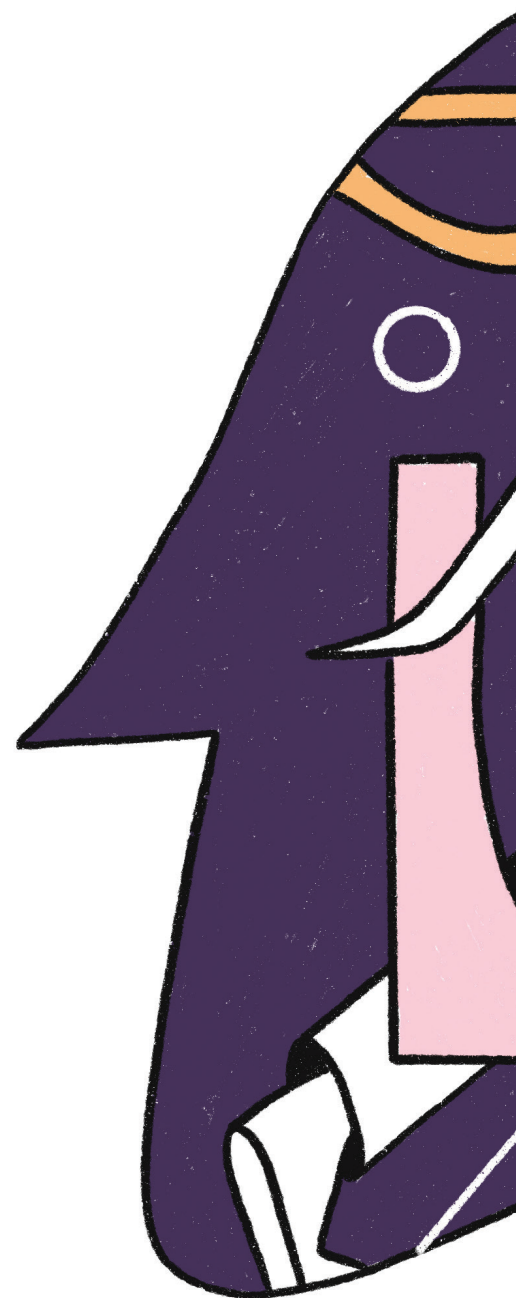
**Y**OU are biased. So am I. We all discriminate. It is both a source of concern and comfort that we don't necessarily do so deliberately and that our prejudices aren't always wilful.

If societies are to truly confront the pernicious effects of racism and prejudice, the importance of examining these biases and how they become etched into the brain is becoming increasingly clear. The death of George Floyd under the knee of a police officer in Minneapolis on 25 May shook the world to attention, but it was no isolated incident. Every day there are stories of people being treated with suspicion – or far worse – based on their skin colour while going about their daily lives.

This is in spite of the fact that, for the past 40 years, opinion polls show a steady decline in racist views in the US, UK and other countries. That has led some researchers to suspect that, as explicit racism has been driven underground, unconscious bias is playing a critical role. This suspicion inspired the creation of the Implicit Association Test, a tool that aims to reveal unconscious biases with a few clicks of the mouse.

Unfortunately, the accuracy and reliability of this widely celebrated test isn't what it once seemed. Pinning down the nature and extent of hidden bias is proving to be extraordinarily complicated. Eradicating it is far from straightforward, too – and it turns out that some efforts to do so may further entrench the very prejudices they are meant to uproot. But we are making progress, not least in understanding the processes in our brains that perpetuate bias – and what we can do to change them.

What exactly is unconscious or implicit bias? In psychological research, the label “implicit” refers to processes that aren't direct, deliberate or intentional self-assessments. When we can't retrieve a memory explicitly, we might still behave in a way that is shaped by our past experiences, for instance. The conscious mind governs deliberate actions, rational thoughts and active learning, while the unconscious carries on with processes that occur automatically or aren't available to introspection. The unconscious is a busy place: the brain is capable of processing approximately 11 million bits of information every second,



## 879

**drug-related arrests take place per 100,000 black people in the US, compared with 332 per 100,000 white people**

Source: US National Survey on Drug Use and Health



but our conscious mind can handle only 40 to 50 of them.

As all of this information comes in, our brains categorise it without our deliberate attention. When we process information on a more superficial level – when we are in a hurry, tired or distracted, for example – we are more likely to rely on existing templates. Occasionally, such cognitive shortcuts can be useful, such as when we need to decide something quickly. But they can also be problematic, especially if these shortcuts were formed based on mistakes, misinterpretations, stereotypes or other biased information. When we use them,

# 77

**black men per 100,000 will be killed by police in the US, based on current trends, compared with 33 white men per 100,000**

Source: PNAS, 2019

we may then be relying on and reinforcing these very mistakes and biases. When that happens with people in positions of power and authority, it can have far-reaching consequences, from discriminatory hiring practices to poorer healthcare treatment or prejudice in the legal system.

The idea that we could pin down and study implicit bias was first hinted at in 1995 when social psychologist Anthony Greenwald, then at Harvard University, and his colleagues invented the Implicit Association Test (IAT) to measure the strength of links between different concepts and words. For instance, participants would be shown black or white faces and asked to pair them with descriptors such as angry, clever, good and bad (see “How the bias test works”, page 40). This was adapted for the web in 1998 by Greenwald and fellow Harvard psychologist Mahzarin Banaji.

There have since been several adaptations of the test, measuring views on race, body type, gender and even names. The array of applications and easy online access have amplified the test’s appeal. It is hard to

THOMAS KULLER

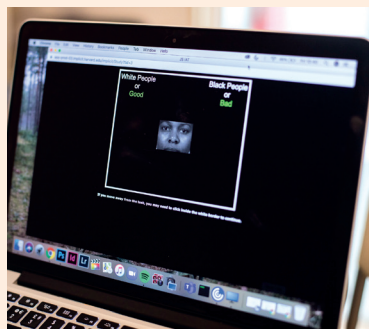


# How the bias test works

The Implicit Association Test (IAT) is an online exercise that involves sorting pictures and words as quickly as possible in a series of tasks using the "E" and "I" keys on a keyboard. For instance, in the weight IAT, you might initially click E if the silhouette of a larger person comes up, and I for a thin one. In the next task, you then sort words with good and bad connotations. Later, you are asked to sort good words and thin silhouettes with one key, bad words and larger silhouettes with another. Then the association is switched.

After completing several sorting tasks, you are given your results in the form of a statement, such as: "Your response suggests a slight automatic preference for fat people over thin people." You can then click through to a page that explains this result: "[O]ne has an implicit preference for Thin people relative to Fat people if they are faster to categorize words when Thin people and Good share a response key relative to when Fat people and Good share a response key."

Through its website, the Project Implicit research group currently offers 15 versions of the test, including on gender, religion, age, skin tone, race, disability and sexuality.



TIM BOBBY

overstate just how influential it has been in both academic research and the public understanding of implicit bias. In his 2005 book *Blink: The power of thinking without thinking*, journalist Malcolm Gladwell summed up the prevailing view: "The IAT

**38**  
per cent

of people from ethnic minorities in the UK report being wrongly suspected of shoplifting, compared with 14 per cent of white people

Source: 2018 Guardian attitudes survey

is more than just an abstract measure of attitudes. It's a powerful predictor of how we act in certain kinds of spontaneous situations."

Yet for all this, its results are inconsistent and hard to reproduce. Many studies have challenged the idea that the IAT reveals only unconscious processes. The reliability of results also appears to decline the more times you take it in a sitting.

What the IAT really measures is reaction time, based on the assumption that the speed with which we make associations reflects underlying mental processes. But everything from reflexes and physical ability to whether the user is distracted can influence this. Several studies have now shown that, for individuals, carrying an implicit attitude is only weakly linked to biased behaviour in the real world.

Part of the problem may be with how the test is used. Neuroscientist Calvin Lai at Washington University in St Louis, Missouri, studies implicit bias and is on the executive committee of Project Implicit, the non-profit research collaboration that studies implicit social cognition and examines the data gathered using the different versions of the IAT. He and others admit that the test is imperfect, but stress that it isn't intended to be a one-off measure. "IAT results should be used as an educational experience for

Anonymised hiring practices can help reduce the influence of bias against people with minority backgrounds



MORSA IMAGES/GETTY IMAGES

self-reflection but should not be treated as a tool for diagnosing one's self or others," he says. "A single administration of the race IAT tells you as much about your enduring racial attitudes as a single measure of your blood pressure tells you about your blood pressure over time: not very much."

## The nature of bias

But aggregated IAT results do tell us something about the nature of unconscious bias within societies. Information from Project Implicit reveals that, of the 630,000 people around the world who have taken a version of the IAT that examines associations between gender and science-related abilities, more than two-thirds correlate males more strongly with science roles and females more strongly with humanities, for instance. Test results from more than 1.8 million people

## Ways to tackle your prejudice

We are still getting to grips with the most effective ways to identify and address bias. What is clear is that it is a difficult task that requires concerted, consistent effort. But there are strategies that make a difference.

A first step is to make biases visible. This can include taking the Implicit Association Test to raise awareness, but this needs to be complemented by active reflection – including recognising triggers for bias and examining how our life experiences have shaped our biases.

Research has shown that using blind or anonymised hiring practices may help weaken biases that can limit opportunities for women and minority groups. One study found that using blind auditions increased the likelihood that women musicians were hired by an orchestra by up to 46 per cent. Research in France, Germany, Sweden and the Netherlands has showed that removing names from applications increases the likelihood that candidates from minority groups will be invited to interview.

We can tackle generalised assumptions

by being clear that a particular attribute is associated with an individual rather than their whole group, for example “This boy is good at maths”. This approach can help to diminish stereotypes and the pressure to conform to them.

Taking our time with important decisions can also help us avoid cognitive shortcuts that perpetuate bias. When this isn’t possible, rehearsing reactions to high stress situations can help prevent biased snap decisions, research with police has shown.

Finding ways to identify with members of different groups by forging links with your own sense of self can diminish bias. In one study, nurses from diverse ethnicities who were shown videos of white or black patients in pain recommended the same amount of pain relief regardless of the patient’s race if first asked to imagine how they felt. When not prompted this way, the nurses suggested more pain relief for white patients. Metaphorically stepping into someone else’s shoes can have a big impact.



in the US showed that in geographic areas where white residents show higher implicit race bias measured by a version of the IAT, there is also greater use of force by the police against black people.

Unfortunately, the IAT is still widely perceived as a diagnostic tool. Most anti-bias courses in the US and UK begin with the test, then give the results as a score that is seldom followed up by a deeper explanation. Occasionally, training programmes give examples illustrating the impact of unconscious bias and tips for how to reduce this influence (see “Ways to tackle your prejudice”, above right).

Yet even with this kind of guidance, bias training is no magic wand that will cure individuals of their prejudices. It doesn’t seem to have a lasting impact on attitudes around diversity within corporations, for example. And while it appears to help reduce

discriminatory behaviour by individuals for up to two weeks after attending, there is no evidence it leads to long-term change. Some kinds of training may even reinforce stereotypes, particularly if the participants are distracted or rushed.

That isn’t to say that we are without options. Advances in brain scanning techniques have helped reveal the neural

# 41

**black women die each year during pregnancy in the US for every 100,000 women who are pregnant, compared with 13 white women and 30 Native American women**

Source: US Centers for Disease Control and Prevention

underpinnings of our biases and in particular how prejudices about other groups of people activate brain areas associated with threat and fear (see “The roots of racism”, page 43). In an influential 2005 study, Mary Wheeler and Susan Fiske at Princeton University asked white volunteers who were in an MRI scanner to perform tasks while looking at black or white faces. They found that when the task involved thinking of the person whose face they saw as part of an out-group, rather than as an individual, the participants showed increased activity in the amygdala, the part of the brain that governs our threat response. Other brain scanning studies show greater activity in the amygdala when people view others from different ethnic backgrounds to their own.

Skin colour isn’t the only way our brains can automatically categorise people. Our response to different accents may be

similar. In 2014, Patricia Bestelmeyer at Bangor University in the UK and her colleagues found that when people heard accents similar to their own, there was increased activity in brain areas associated with positive emotional response; the opposite was true for different accents. “There is an increasing perception of the importance or relevance of those accents that are similar to ours,” she says.

Yet the imaging revolution in bias research has also demonstrated that our brains can change with experience and environmental influences. In 2013, Eva Telzer, then at the University of Illinois, and her colleagues conducted a study of 49 children and adolescents born in Asia, Europe and the US. They showed that the difference in amygdala activity in response to faces from different races wasn’t innate, but developed over a period of time.

This landmark study quashes any suggestion that we are somehow born prejudiced. What’s more, Telzer and her team found that study participants with a more

**15**  
per cent

**of job applications from ethnic minority candidates in the UK received a positive response compared with 24 per cent of those from a majority group – despite both sets of CVs having identical qualifications**

Source: GEMM Project, Centre for Social Investigation, Nuffield College, University of Oxford

**The diversity of peer groups can influence who we perceive as threatening**

diverse set of peers had less of a heightened threat response in the brain when shown faces from other racial groups. That suggests simply having more contact with people from different groups can reduce the importance of race in how we respond to people and that we can change our biases.

This wasn’t always a given. In social psychology there was a long-standing assumption that traces of past experiences linger on whether we want them to or not. But we now know that unconscious bias isn’t as stable as previously believed. Our biases are shaped by how we are brought up, what we see around us and the media we are exposed to. Knowing we can change their influence also means we can no longer shrug them off as beyond our control.

One day we may even have a tool that helps us to reliably measure them. “There is ongoing research to develop longer or more sophisticated versions of the IAT or other implicit measures that are reliable enough for diagnosis,” says Lai. Unfortunately, none are yet ready for public use.

We needn’t wait for new tools to assess the harms of bias, though. “Your best bet for understanding inequities in your organisation is collecting data about inequities within your organisation, not taking the IAT,” says Lai.

Even as efforts are under way to better measure the influence of unconscious bias, a growing number of researchers argue that we actually need to simplify this debate – to drive home that bias is bias, and whether it is unconscious or overt, whether individual prejudices shape social institutions or are shaped by them, they can cause irreparable damage. Unconscious bias is easier to ignore, but it cannot excuse discriminatory behaviour. It is important to remember that even if we cannot precisely measure our biases just yet, we can still overcome them. ■



THOMAS BARWICKGETTY IMAGES

SIMON SONGHURST



Pragma Agarwal is the author of *Sway: Unravelling unconscious bias* (Bloomsbury). Follow her @DrPragmaAgarwal