# Is unconscious bias training still worthwhile?

*Training can raise people's awareness of their unconscious biases, but evidence shows that training alone is not effective in changing behaviour. The UK government has already decided to discontinue this kind of programme in its various departments.* ***Frederick Herbert*** *writes that while it is generally accepted that awareness is not a sufficient condition for behavioural change, it is usually necessary. He argues that unconscious bias training can be re-thought of as a foundation upon which other interventions can build.*

---

On 17 December the government released a written statement announcing that "unconscious bias training would be phased out in [government] departments" (Lopez, Julia, 2020). To many people, the news may have immediately smacked of regressive decision-making and a lack of interest in the experiences of diverse employees. In reality, as the evidence for UBT is generally lacking, it feels more like a missed opportunity than a failure.

Let's start with the easy bit. Unconscious bias is important. There is robust evidence that automatic mental associations about social groups impact our perceptions and decision-making and that people tend to hold more negative associations toward a variety of historically oppressed groups. This can have serious impacts on people's lives. For instance, systematic reviews show that bias amongst medical healthcare professionals is common and is associated with a reduced quality of care (FitzGerald & Hurst, 2017). Similarly, there is experimental evidence that skin colour primes decision-making such that "participants shot armed Blacks more quickly than armed Whites and decided not to shoot unarmed Whites more quickly than unarmed Blacks" (Correll et al., 2006). Clearly, it is not hyperbole to say that lives are on the line.

In fact, the only disputable part of this debate is how large a proportion of the many examples of prejudice in modern society result from these negative implicit associations, as opposed to rank and deliberate discrimination. Either way, in a world where submitting a resume with a white sounding name results in 50% more call-backs for interviews than resumes with African-American names, it is clear there is plenty of work to be done to overcome these effects (Bertrand & Mullainathan, 2004).

Against this backdrop, the government's decision to phase out unconscious bias training (UBT) from the civil service may seem unconscionable to some. However, the patchwork evidence for UBT does warrant genuine concerns about its value, given the time-related costs of providing training for the hundreds of thousands of civil servants. To evaluate this decision, it's important to first understand the evidence.

**The evidence**

So, what does the evidence say? Is unconscious bias training, as a one-off standalone training program, a useful intervention? Answering this requires us to be clear on what results we're expecting from unconscious bias training. In general, four possible benefits have been investigated.

– To increase knowledge of the existence and impact of negative implicit associations
– To reduce negative implicit associations.
– To change explicit attitudes
– To change behaviours in such a way that those who attend exhibit significant reductions in prejudice.

To give a fair and balanced answer it's worth turning to each of these in turn:

Knowledge is the simplest outcome to achieve. It is therefore unsurprising that the evidence here is strongest. A recent large-scale meta-analysis of all forms of diversity training showed that awareness tended to significantly increase amongst participants across the studies reviewed (Bezrukova et al., 2016). The capacity for interventions to increase knowledge of implicit prejudicial associations was also explicitly acknowledged within the evidence summary the government decision was based on (The Behavioural Insights Team, 2020). In short, UBT does seem to teach people about negative implicit bias.

Date originally posted: 2021-03-24

Permalink: https://blogs.lse.ac.uk/businessreview/2021/03/24/is-unconscious-bias-training-still-worthwhile/

Blog homepage: https://blogs.lse.ac.uk/businessreview/

Reducing negative implicit associations is perhaps the most common metric within studies. This makes sense as unconscious bias is itself an implicit phenomenon. Some studies have shown positive effects where implicit bias towards women in STEM have bene reduced (Jackson et al., 2014). However, others have shown that there may even be unintentional backfire effects which strengthen negative associations (Payne et al., 2002). Most importantly, a large-scale meta-analysis of interventions to change implicit associations found that effect sizes were generally small or not distinguishable from zero (Forscher et al., 2019). That is, any changes from training were negligible. A systematic review focused specifically on reducing implicit prejudice reached similar findings (FitzGerald et al., 2019). Forscher's analysis also revealed that in more robust studies "effect sizes tended to be smaller than their corresponding overall meta-analytic estimates". This suggests that even the small effect sizes found are probably overestimated due to various forms of research bias. While these are unimpressive results, to some extent it is also unsurprising. The most common conceptualisation of implicit associations is that they form as a result of a lifetime of exposure to associations and are unconscious (Rudman, 2004). Within this paradigm a one-off intervention is unlikely to make a significant impact on one's associations when weighed against a lifetime of exposure to prejudice. However, it should be noted that other evidence suggests implicit associations can be more unstable and changeable than explicit attitudes (Gawronski et al., 2017). In summary, there is no strong evidence for the effectiveness of one-off UBT interventions to reduce implicit prejudice.

Explicit attitudes are a simpler way to measure the impact of an intervention. Rather than a test of associations, this relies on self-reported attitudes. However given the sensitivity of the topic, the accuracy of self-reports is at best highly questionable due to the strong likelihood of a social desirability bias (Forscher et al., 2019; Jackson et al., 2014). This problem makes it hard to interpret study results, as social desirability bias could conceivably be skewing both control and experimental group responses. In fact, scepticism about the validity of this measure was one reason implicit measures began to gain prominence. Whatever the cause, the meta-analysis results are once again unimpressive. Explicit attitudinal change is not reliably found, and when there are positive changes, the results are even smaller than for implicit measures (Forscher et al., 2019).

Finally changing behaviour, which could be thought of as the gold standard metric, as it is the pathway likely to have the most tangible impact on people's lives. As behaviour change generally follows and is smaller than attitudinal change (Webb & Sheeran, 2006) it may be expected that any changes would be smaller still than attitudinal measures. However, there is some evidence that diversity training that teaches people about discrete actions they can do differently is more successful than when the focus is on attitudes (Bezrukova et al., 2016). Unfortunately, this evidence is focused on diversity training in general, rather than implicit prejudice, so it is of interest only tangentially. When looking at UBT interventions alone, once again, there was no significant evidence of behaviour change (Forscher et al., 2019).

To summarise, it seems only knowledge was reliably shifted by UBT interventions. Other metrics had small to negligible effects in the short term, and worse results in the long term. There is also evidence that poor research practices have inflated results in the existing literature, and that backfire effects may occur in some instances.

Given this evidence from recent meta-analyses, policy makers and practitioners may be excused for having been led down the garden path and back again by the patchwork scientific evidence and the ever-present crises of bias and inflated results within social sciences (Shrout & Rodgers, 2018).

You may be surprised to hear then that I still don't think scrapping UBT was the correct decision. Let me explain why.

**Is knowledge enough?**

The first key point of challenge is to question whether the real problem is expectations rather than outcomes. Personally, I believe there is a strong argument that increasing awareness is still valuable.

Date originally posted: 2021-03-24

Permalink: https://blogs.lse.ac.uk/businessreview/2021/03/24/is-unconscious-bias-training-still-worthwhile/

Blog homepage: https://blogs.lse.ac.uk/businessreview/

Within the academic literature it is generally accepted that while awareness is not a sufficient condition for behaviour change, it is usually necessary (Chang et al., 2019; Devine et al., 2012; Equality and Human Rights Commission, 2018). This view fits with the widely used COM-B model of behaviour. In this framework, behaviour is a product of capability, motivation, and opportunity. 'Capability is defined as the individual's psychological and physical capacity to engage in the activity concerned. It includes having the necessary knowledge and skills' (Michie et al., 2011). Behaviour change frameworks, such as the Theoretical Domains Framework (TDF), also emphasise that knowledge is a key tool in behaviour change (Atkins et al., 2017). This is not a revolutionary thought, but it is important.

From this perspective UBT can be re-thought of as a seed or foundation upon which other interventions can build. As Fitzgerald et al. state, even if UBT is not effective at changing behaviour as a standalone intervention, it may still be important as "we would expect a virtuous person who finds discrimination based on race abhorrent to be disturbed to discover that she automatically associates a historically oppressed race that still suffers discrimination with negative qualities" (FitzGerald et al., 2019). If UBT allows people who would never self-identify as prejudiced to realise that their bias can result in problematic decision-making, it is doing something. As post-apartheid South Africa discovered, albeit in a far more extreme circumstance, there is a power in truth, however unsavoury that truth may be to hear (Allan & Allan, 2000).

For those reviewing the effectiveness of UBT, awareness of implicit associations and unconscious bias may seem like a measly return. But how many civil servants in the country have never realised their implicit negative associations are likely to be contributing to systemic racism or sexism?

**Indirect effects**

The second factor to consider is the potential for indirect benefits of UBT and indirect costs from the decision to reverse it. Firstly, in the written ministerial report I found no evidence of a consultation with various stakeholder groups on the decision. The civil service have established and well-organised networks that represent the groups that most commonly experience prejudice in the workplace. If these groups find it valuable to have a forum where the day-to-day bias they face is acknowledged and recognised, then there's also an argument to be made that UBT is helping to create a more inclusive work culture. Notably, in the independent review of race in the workplace commissioned by the Department for Business, Energy, and Industrial Strategy (BEIS), which had a strong consultative aspect, UBT was backed as a worthwhile intervention (BEIS, 2017).

Secondly, UBT may be thought of as a key signalling device that shows managerial commitment to diversity. As might be expected, managerial commitment to a topic has been linked with a workforce's willingness to engage with it (Salas & Cannon-Bowers, 2001). While more evidence is needed to explore how significant an impact managerial signalling might have on behaviours, removing the only mandatory diversity training for civil servants with no direct replacement is certainly not fantastic optics.

To be fair, the government's statement does mention that a new strategy will be released, and that this strategy will be evidence-led and will take on core learnings from the various meta-analyses on diversity training. For instance, the statement mentions a shift to a more integrated long-term approach, which will move "inclusion and diversity into mainstream core training and leadership modules in a manner which facilitates positive behaviour change" (Lopez, Julia, 2020). If diversity training is embedded and integrated rather than buried, this will be a good thing. But it doesn't make much sense to create a vacuum by first scrapping UBT and then delaying the announcement for what follows.

**Take the lead**

Date originally posted: 2021-03-24

Permalink: https://blogs.lse.ac.uk/businessreview/2021/03/24/is-unconscious-bias-training-still-worthwhile/

Blog homepage: https://blogs.lse.ac.uk/businessreview/

Most importantly, instead of scrapping UBT the government could have made itself a leader and have truly striven to understand what works in the most robust possible way. I spent many hours going through the evidence, and my main takeaway wasn't that the evidence for the effectiveness of UBT was poor. Instead, it was that, as-ever, interventions in externally valid settings are incredibly complex to understand and measure. Given the diversity of interventions, outcomes, participants and research methodologies in this field, an average treatment effect is only ever going to be a small part of the story. For instance, in the most compelling systematic review of interventions focusing specifically on reducing implicit prejudice, only six studies from the UK were considered robust enough to include. These six studies focused on different types of bias (age, religion, obesity, and race). They used different types of interventions, including imagining positive contact, educational films, discussing commonalities, embodiment in black avatars, and even a single oral dose of propranolol. In many of these the recipients of the interventions were primarily young female psychology students (FitzGerald et al., 2019). So, what does this review tell us about the effectiveness of the UBT currently used within the civil service? Not very much at all.

While they do tell us that on average bias reduction techniques don't work that well, the meta-analyses and reviews I looked at were also clear that there is significant variability in outcomes based on the specifics of the intervention, and that there were some common factors amongst successful interventions. For instance, in successful treatments, participants were highly involved and able to identify with the people used in example scenarios (FitzGerald et al., 2019). All this begs the question, rather than relying on this patchwork literature, why not set up an internal civil service trial to test the effectiveness of the current training program? Better yet, why not develop a multi-armed trial of UBT including a few adapted versions of the current training based on uncovered suggestions of best practice? This would be the best and most valid way to understand what works within the civil service context and would also help strengthen the evidence base more generally. It would have also been a fantastic signal that the government cares about leading the way in uncovering how to reduce bias and prejudice in the workplace.

**Conclusions**

In the end, it's hard to take too strong a stance on this topic. The government is right that there is little compelling evidence for the effectiveness of UBT. However, for now, the decision to scrap UBT seems like a missed opportunity for the government to take a lead by working with workplace networks and running trials to understand what works when it comes to tackling discrimination in the workplace. The impact of bias remains pervasive. Let's hope the forthcoming strategy pulls out all the stops.

*Notes:*

- *The post gives the views of its authors, not the position of LSE Business Review or the London School of Economics.*
- *Featured image by Jr Korpa on Unsplash*
- *When you leave a comment, you're agreeing to our Comment Policy*

Date originally posted: 2021-03-24

Permalink: https://blogs.lse.ac.uk/businessreview/2021/03/24/is-unconscious-bias-training-still-worthwhile/

Blog homepage: https://blogs.lse.ac.uk/businessreview/