

Opinion

# Implicit bias reflects systemic racism

B. Keith Payne<sup>1,\*</sup> and Jason W. Hannay<sup>1</sup>

**Implicit bias refers to automatically evoked mental associations about social groups. The idea has been influential across the social sciences as a way to explain persistent racial disparities amid changing self-report attitudes. Most research has treated implicit bias as an individual attitude. However, evidence suggests that it is neither a stable individual difference nor a strong correlate of individual behavior. Moreover, the individual-focused approach can lead researchers to neglect systemic racism as a cause of persistent disparities. We argue that implicit bias can be considered a cognitive reflection of systemic racism in the environment. In this view, implicit bias is an ongoing set of associations based on inequalities and stereotypes in the environment. As such, implicit bias changes when contexts change.**

## Persistent disparities amid changing attitudes

Protests against racism, sparked in part by the killing of George Floyd by Minneapolis police, may be the largest social movement in US history. As of this writing, tens of millions have participated [1]. This massive movement has increased recognition of systemic racism, long limited to mostly academic discourse. With the wider use of this concept, more people are realizing that discrimination does not result only from prejudiced individuals, but also from structures or procedures that systematically disadvantage marginalized groups. In this paper we argue that implicit racial bias, long considered an individual attitude, may be a better indicator of systemic racism than individual prejudice.

Scholarship focusing on individual prejudice has generally supported a narrative of progress. Survey data suggest that racial attitudes have gradually trended more egalitarian for decades. For example, between 1990 and 2017 the percentage of White respondents saying that Black people work less hard than White people declined from 65% to 31%, and the proportion saying that White people are more intelligent than Black people declined from 56% to 20% [2]. Many people still endorse these stereotypes, but the trend is often interpreted as progress.

Scholarship on systemic racism supports a less optimistic view. Segregation in housing and education declined in the 1960s and 1970s, but have remained stable since then [3]. Large racial gaps in chronic diseases and life expectancy have remained steady since at least the 1980s [4]. Disparities in income and wealth have grown since 1979 (<https://files.epi.org/pdf/101972.pdf>).

Understanding the persistence of systemic racism in the context of changing individual attitudes has set the agenda for a great deal of research across the social sciences. One of the most important contributions from psychological science is the concept of implicit bias. Implicit bias refers to positive or negative mental associations cued spontaneously by social groups. It is measured using cognitive tasks that test how those associations facilitate, interfere with, or otherwise bias task performance [5,6]. Many studies suggest that implicit bias is widespread, even among people who explicitly endorse egalitarian attitudes [7,8]. Psychologists have often looked to implicit bias to explain persistent discrimination and disparities in the face of declining explicit

## Highlights

Implicit bias has typically been understood as an individual attitude. We challenge this assumption by showing that implicit bias is neither a stable individual difference nor a robust predictor of individual behavior. Instead, we argue that implicit bias reflects the mind's ongoing predictions based on regularities in the environment.

This context-based view of implicit bias suggests that the average level of bias in an environment, such as a workplace, city, or country, is diagnostic of the systemic racism in that environment.

Implicit bias may serve as a marker of – and a driver of – systemic inequalities.

Although efforts to reduce implicit bias may be effective in the short term, longer term change is more likely to come from directly changing systems, policies, and processes that sustain inequalities.

<sup>1</sup>Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

\*Correspondence: [payne@unc.edu](mailto:payne@unc.edu) (B.K. Payne).



prejudice. Although it has often been treated as an individual attitude, we argue that implicit bias can best be seen as a cognitive reflection of systemic racism in the environment. In this opinion article, we argue that implicit bias is a byproduct of a mind that draws inferences based on statistical regularities, when that mind is immersed in an environment of systemic racism.

### **Implicit bias as an individual attitude**

Much early research treated implicit bias as an individual attitude internalized through early social learning [9–12]. Once learned, implicit bias in this view is considered difficult to change. For example, when interventions fail to show lasting changes in implicit bias, those failures are often interpreted as evidence for the tenacity of individual attitudes [13,14]. According to the individual attitude perspective, individual differences in implicit bias are important because individuals who have high bias are more likely to discriminate than those with low bias [15]. Reducing discrimination thus requires understanding who has high levels of implicit bias, and how their attitudes can be changed [16]. In summary, the individual view assumes that implicit biases are stable attitudes that differ from one person to another, and these individual differences are associated with biased behavior.

The individual attitude view has been questioned because accumulating evidence suggests that it is neither strongly associated with individual behavior nor stable over time. Meta-analytic summaries of associations with behavior range from  $r = 0.24$  [17] to  $r = 0.14$  [18] for the Implicit Association Test (IAT) and to  $r = 0.28$  for sequential priming measures [19]. A correlation of 0.28 implies that only 8% of variance in behavior measures can be explained by implicit test scores. Some researchers suggest that small correlations in this range are typical for behavioral sciences, and can be important even if numerically small [15,20]. However, even if an individual difference is weakly related to behavior, most researchers agree it should be stable over time.

Test–retest stability for implicit bias is quite low. In a meta-analysis, the average test–retest stability for the IAT was  $r = 0.42$  [21]. Temporal stability appears to be similar for the Affect Misattribution Procedure [21,22]. Both tests show much higher internal consistency, which suggests that implicit biases can be reliably measured at a given moment, but they are unstable over time [19,21]. Although there are no objective standards for how reliable is reliable enough, researchers often consider a test–retest correlation of 0.70 or higher to be acceptable as a rule of thumb [23]. A test–retest correlation of 0.42 suggests that only 18% of the variance in test scores can be explained by the same participants' previous scores.

However, reliability and validity look dramatically different when examining average levels of implicit bias rather than individual differences. The test–retest correlation for state-level IAT scores is  $r = 0.76$  from one year to the next, and  $r = 0.69$  across a decade [24]. Average levels of implicit bias are also strongly correlated with racial disparities in consequential outcomes. For example, countries with stronger average associations between males and science had greater gender-based achievement gaps in eighth grade science ( $\beta = 0.55$ ) and math scores ( $\beta = 0.67$ ) [25]. Metropolitan regions in the USA with higher levels of implicit race bias have greater racial disparities in police shootings,  $\beta = 0.39$  [26]. County-level implicit bias is associated with health disparities [27,28]. State-level implicit bias is strongly associated with the frequency of internet searches for racial slurs,  $r = 0.78$  [29]. This type of evidence led us to shift from thinking about implicit bias as a feature of individuals to thinking about it as a feature of contexts.

### **The Bias of Crowds**

The Bias of Crowds model [24] reconceptualizes implicit bias in terms of contexts and systems. Most theories agree that implicit bias reflects the accessibility of mental associations [30].

Concept accessibility refers to the readiness with which information (including attitudes and stereotypes) can be retrieved and used in cognitive processing [31,32]. Like other theories, the model assumes that when we use one concept, other related concepts and information become more likely to be accessed for use and that this process happens spontaneously and involuntarily. The accessibility of thoughts and evaluations associated with social categories such as Black, White, male, or female is what implicit tests measure.

The accessibility of information in memory tracks statistical regularities in the environment [33]. The frequency with which a word occurs in the language, for example, determines how easily retrieved that word is. The frequency with which two words co-occur determines how likely one associate is to come to mind when the other is presented. That is why *Mississippi* cues *river* more than *rifle*, and why, when it comes to stereotypes in US culture, *Islamic* cues *terrorism* more than *terrarium*.

To be clear, the idea that concept accessibility tracks statistical regularities does not imply that stereotypes are accurate or that prejudices are based in reality. Most White Americans' experience with Black Americans is largely indirect, through media that are themselves riddled with stereotypic biases [34,35]. The 'rubbish in, rubbish out' principle applies: a cognitive system that forms associations based on statistical regularities will be systematically biased, so long as it operates in an environment that is characterized by pre-existing stereotypes and inequalities [36].

It is uncontroversial that accessibility can vary both chronically [37] and situationally [38]. For example, one person may have a link between the category of Black Americans and a negative stereotype that is more chronically accessible than that same link is for another person. This notion of a chronically accessible link maps onto what most theorists mean by implicit attitudes: relatively stable dispositions based on learned associations.

Accessibility can also vary temporarily, as a function of the context. Two decades of experiments demonstrate that average levels of implicit bias are malleable in response to manipulations of the situation. For example, average racial bias scores can be reduced by activating counter-stereotypical associations, interacting with an African American experimenter, reading a counter-stereotypical story, or thinking about positive exemplars like Martin Luther King Jr or Denzel Washington [39,40]. Scores can be increased by reading a story, listening to music, or thinking about examples that confirm racial stereotypes [41,42] (for reviews see [43,44]).

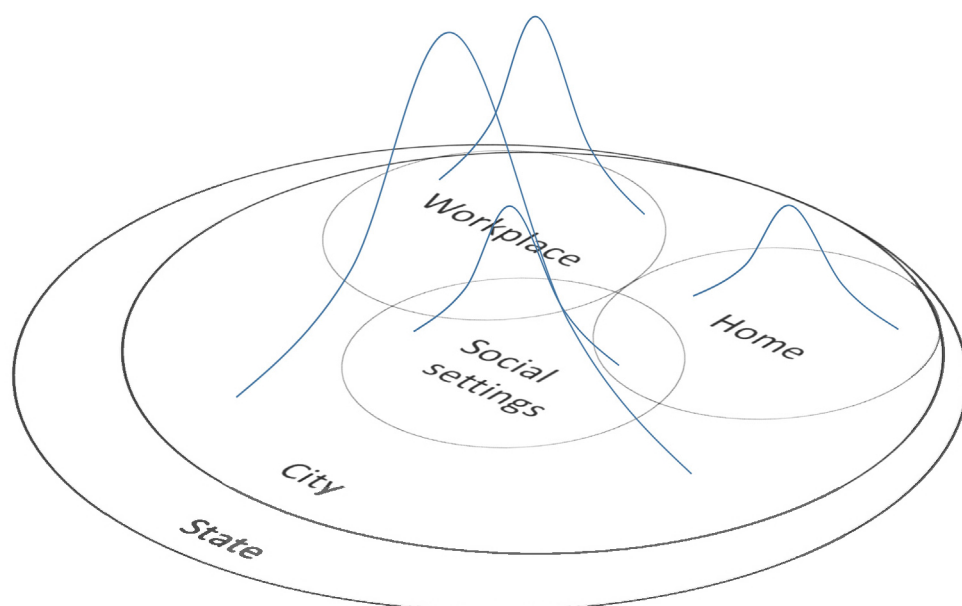
More recent studies suggest that implicit evaluations can be eliminated or even reversed based on changes of context. For example, taking the perspective of a defense attorney led to more positive (or at least less negative) implicit evaluations of Adolf Hitler [45–47]. Other studies suggest that implicit evaluations can be reversed by leading participants to reinterpret information about a person [48]. These studies suggest that changing what information is situationally available can alter implicit bias scores.

These demonstrations of situation-based malleability are well-established, but researchers disagree about what they mean for the nature of implicit bias. Some researchers interpret context effects as changing only test performance, but not affecting 'true' implicit bias, understood as an individual attitude [49]. Others argue that implicit bias is a stable trait-like construct, and that context effects or temporal fluctuations reflect only measurement error [50,51]. We argue, by contrast, that effects of the context on aggregated scores are interesting and important for understanding the nature of implicit bias itself.

How do unstable individual scores add up to stable and valid measures of the social context? Research on ‘the wisdom of crowds’ finds that, for many kinds of questions, the collective judgment of a group tends to be closer to the true answer than any one individual’s answer [52,53]. Crowds are ‘wise’ because each individual is likely to have partial true knowledge as well as errors that are largely random. When independent judgments are averaged, the random variations are aggregated away, leaving the true knowledge to emerge as the central tendency of the distribution. In the same way, the statistical benefits of aggregation allow average implicit bias scores to produce an accurate estimate of the most widely shared implicit biases in that context.

Contexts that cue biases are partially overlapping (Figure 1). For example, one’s work and social environment may partially overlap, because people often socialize with co-workers. Both likely fall mostly within the city, state, and country in which one lives. Each context has a theoretical distribution of implicit bias. Implicit test scores in a sample of participants can therefore potentially measure implicit bias at multiple levels at once. When aggregated, average bias scores will reflect the accessible knowledge that is shared in common across that context. Averaging across cities, for example, will estimate the most widely accessible biases in each city, whereas averaging across universities will estimate the most widely accessible biases in each university. Contexts can also be virtual, such as social media networks or mass media markets. For this reason, aggregated implicit bias scores can be considered a measure of systemic racism, defined at any given level of the ‘system’.

The Bias of Crowds model is consistent with other frameworks that emphasize the role of context. Research on bias as a regional construct [54] has treated cities, counties, or states, as the units of analysis. By establishing the construct validity of implicit and explicit biases at these regional levels, this research focuses on regions rather than individuals. It makes no claims, however, about psychological differences between person and regional level measures. The



Trends in Cognitive Sciences

Figure 1. Hypothetical distributions of implicit bias in social contexts. Social contexts are partially overlapping. Each has a distribution of implicit bias. Aggregating implicit bias scores across a given context can estimate the implicit biases shared most widely within that context, even if the biases of individuals in the context change over time.

Prejudice in Places model [55,56] also highlights the role of contexts. It does not distinguish between implicit and explicit bias, but shares the emphasis on the role of social environments such as norms, policies, and procedures, in creating advantages and disadvantages. From the Bias of Crowds perspective, implicit bias operates differently than explicit bias because implicit bias is less stable and trait-like, and hence more easily tracks changes in contexts. Nevertheless, all of these frameworks redirect attention to systemic sources of bias and systemic solutions.

Nearly any measure can be examined as a function of both persons and contexts. However, the effects of aggregating person variables to measure contexts has different effects, depending on the properties of individual-level observations. For highly stable traits, properties of the aggregate measure can be fully reduced to properties of the individuals. Aggregation has a different effect for characteristics that are highly malleable, such as implicit bias. Stable aggregate bias is an emergent property in that it cannot be reduced to stable individual scores. A city might have the same mean year after year, although the rank orders of the individuals composing that mean vary dramatically from one day to the next. The stability that emerges from aggregating implicit bias scores resembles the constant surface tension of a balloon, produced by billions of randomly moving air molecules.

This *emergent stability* is interesting because the mean level of implicit bias in a context can reveal information about the context that may not be reflected in any individual's stable attitude. Consider the way that 'the wave' moves through a crowded sports stadium. The wave travels through the crowd with a stable velocity of about 20 seats per second [57]. Once the wave has begun in one section, we can accurately predict when a given section of fans will stand up. Although the wave is composed of individuals, it does not matter much which individuals make up the crowd. The Bias of Crowds model suggests that implicit bias, like the wave, is a social phenomenon that passes through individuals rather than only residing inside them. Importantly, this implies that one could largely exchange individuals, but the average of the crowd would remain the same.

Although many authors (including ourselves) acknowledge that implicit bias scores are influenced by both person and context factors [58,59], the Bias of Crowds takes a further step that other perspectives have not. We argue that to the extent that contexts lead to reliably different aggregate scores, the best interpretation of those aggregate scores is about features of those contexts. When a sample of people displays bias on an implicit test, the most apt conclusion is not that they are biased people, but that they are in a biased context. This interpretation raises questions about what in the context cues biases, as opposed to what kind of person holds biases. Once implicit bias is reconceptualized as a property of contexts, it changes both the ways that implicit bias is measured, and the nature of the psychological phenomenon we are trying to explain, as we elaborate next.

### Implications for measurement

A practical implication of the model is that implicit bias can be measured more reliably by aggregating scores. We argue that this is an advantage that researchers should utilize whenever possible. It is good science to measure at the level where one can measure reliably.

One critique of the Bias of Crowds argues that the low stability and small associations with behavior result only from error variance in the measurement of stable individual attitudes [60]. However, it would be a mistake to equate situational malleability with measurement error, because that malleability reflects, in part, meaningful variation across contexts. Error variance refers to variance that is unrelated to the construct of interest. How researchers define error variance therefore depends on the researchers' interests.

From a person-focused perspective, changes across time and contexts are error variance. However, from a context-focused view they are the true score of interest (for a detailed discussion see [61]). Separating these sources of variance requires longitudinal and/or multi-level study designs that have rarely been used in studying implicit bias. Such designs can separate true score and error variance at the levels of persons, contexts, time, and so on.

### Implications for the psychology of implicit bias

The person versus situation debate is common across many topics in psychology. However, the substantive psychological meaning of person and situation factors varies depending on whether the topic is altruism, emotion, obedience, implicit bias, and so on. Person-based theories of implicit bias have focused on explaining attitudes as a function of learning histories [12].

We do not dispute that person effects (chronic accessibility) exist, but we argue that situational accessibility effects have a distinct psychological meaning. Situational accessibility effects highlight the transient activation of concepts. Cognitively, this is more akin to semantic priming rather than attitudes or beliefs. Analyzing implicit bias as a priming effect rather than a static attitude leads to different questions and different explanations. The person-based view leads to explanations of implicit bias in terms of childhood experiences [62] and the qualities that differentiate one person from another [50]. The situationist view prompts entirely different explanations that do not involve childhood experiences, early learning, or slow-learning memory systems. Instead, explanations are about what environmental cues activate bias, how cognitive systems form real-time inferences based on regularities in the environment, and how momentarily accessible concepts influence behavior.

Aggregating scores at the college campus level, for example, affords questions such as what features of the faculty, the student body, or the built campus environment might contribute to different levels of bias [63]? Aggregating at the state or country level afford still other questions. These questions are not only different, but often the questions asked at one level would not make sense at another level (Box 1). The individual-level questions outlined in Box 1 have been studied extensively. The context-level questions are only beginning to be addressed.

### Implications for systemic racism and systemic change

Efforts to reduce implicit bias have often focused on the individual. Guided by the individual attitude model, implicit bias training tries to change attitudes by education or persuasion. The context-centered view of implicit bias suggests that changing contexts will be more effective [56,64].

#### Box 1. Questions afforded by person-level and context-level analyses

Person-level questions:

- How do people learn implicit biases? Why do some people form strong biases while others do not?
- How, why, and under what conditions do implicit attitudes predict individual behavior?
- How can implicit attitudes be changed? What persuasion, education, or training techniques are effective at changing attitudes?

Context-level questions:

- What are the root causes of implicit biases? What historical forces, policy choices, or cultural features make some contexts more biased than others?
- What proximal cues or signals make implicit biases more accessible in some places than others?
- How can institutions or organizations be designed to reduce implicit biases? What policies or procedures are important for making a place less biased?



Many strategies apply to organizations or institutions. For example, one way to change an organizational context is to change local norms, such as workplace expectations that communicate that discriminatory behavior or prejudiced comments are unacceptable. A second way to change contexts is to increase visible representation of minority employees, especially in positions of power. One study, for example, found that universities with more racially diverse faculties displayed lower implicit race bias among the undergraduates [63]. A third way is to revise policies and procedures, such as hiring, employee evaluation, and reward policies that may disadvantage some groups, even if unintentionally. Training efforts should focus not on changing individuals' attitudes, but on learning how policies and procedures can decrease bias.

Other strategies for systemic change involve large scale political policies. The research reviewed above suggests that changing social norms are important for changing both implicit bias and behavior [65,66]. Much of the intense debate over 'political correctness', 'wokeness', and 'cancel culture' are about changing norms and reactions against those changes. Recent protests are another attempt to focus public attention on racial injustice and to motivate change. Policies that reduce inequalities directly, such as legalizing same-sex marriage, are another potent approach. Likewise, policies that directly increase racial equality and participation, such as voting rights legislation and affirmative action policies, are likely to have more impact than changing attitudes.

### Novel predictions from the Bias of Crowds model

The Bias of Crowds model generates novel predictions about the nature of implicit bias in particular (see [Outstanding questions](#)). One novel prediction is that regional differences in implicit bias are caused by features of the context, rather than resulting simply from the clustering of people with similar attitudes. As a result, aggregate levels of implicit bias should be relatively robust to changes in the individuals composing the crowd. For example, the model suggests that city-level implicit bias should have similar stability and predictive validity in cities with high or low rates of population mobility.

A second novel prediction is that the well-documented temporal instability in implicit bias measures does not simply reflect measurement error. Instead, temporal fluctuations reflect meaningful changes as implicit bias passes through individual minds. Thus, for an individual, implicit bias may be more like the stream of thought than a rigid attitude.

A third prediction is that aggregating multiple tests per person should result in relatively little improvement in measurement validity. If temporal instability reflected only measurement error, then aggregating more observations per person should greatly improve predictive validity by creating a better measure of the stable attitude. However, to the extent that implicit bias is itself changeable rather than stable, aggregating more measures per person should have modest effects.

A fourth novel prediction is that implicit bias should not always be rigid (as was assumed in early learning theories) nor should it always be fleeting. Instead, aggregate implicit bias should change at the rate that the social environment changes. Social environments can exist at multiple levels of analysis, which may change at different rates. As we discuss next, recent evidence suggests that when the context changes, aggregate implicit bias changes as well.

In some cases, context change is excruciatingly slow. Laws, institutions, norms, and cultural and historical forces can create patterns of disadvantage that last for generations. Examining an implication of the Bias of Crowds model, one study [67] investigated the long-term legacy of slavery

on modern-day implicit bias. The study found that both counties ( $r = 0.37$ ) and states ( $r = 0.63$ ) with larger enslaved populations in the 1860 census have higher levels of pro-White implicit bias today among White residents. These associations were partially mediated by measures of modern structural inequalities, including residential segregation and racial disparities in economic mobility. Whereas individual levels of bias are unstable over 2 weeks, this study was able to detect the residue of slavery in implicit bias after 160 years. This persistence is difficult to explain if implicit bias is viewed as an individual attitude, but it can be naturally explained by viewing implicit bias as a reflection of biases in social contexts, which are shaped and maintained by structural forces.

In some cases, change comes gradually. A study of IAT scores in the US between 2007 and 2016 found that implicit bias regarding race and sexual orientation have become more egalitarian over time [68]. However, biases regarding age and disability remained stable, while biases regarding body weight became stronger. According to the contextual view, these changes likely reflect gradual cultural changes, which may affect whoever experiences that culture. One source of contextual change is ethnic and racial diversity. One study found that metropolitan areas with greater multi-ethnic diversity displayed lower implicit associations linking Black Americans and weapons [69]. Moreover, metros where diversity increased the most over time exhibited reductions in implicit biases [70,71].

Sometimes, contexts change quickly. One study found that implicit anti-gay bias declined more quickly in states after they legalized same-sex marriage [72]. Another study found that average anti-fat implicit bias increased in the weeks following fat-shaming events in mass media [73]. These findings suggest that widely shared experiences can shift aggregate implicit biases rapidly.

## Concluding remarks

As self-reported racial attitudes have trended gradually more egalitarian for decades, actual racial disparities have hardly changed. The concept of implicit bias was developed to help explain that gap, but most research has treated it as a static individual attitude. Reconceptualizing implicit bias as a cognitive reflection of systemic bias helps move this research forward. Applying this new approach to understand implicit bias raises new questions such as: What features in the environment cue biases? What social structures perpetuate biases over time? How can environments be designed to reduce bias? Answering these questions requires a conceptual shift away from thinking of implicit bias as primarily a rigid attitude. It may be more generative to think of implicit bias as the natural outcome of a mind that generates associations based on statistical regularities, whenever that mind is immersed in an environment of systemic racism.

## Declaration of interests

No interests are declared.

## References

- Chenoweth, E. (2020) The future of nonviolent resistance. *J. Democracy* 31, 69–84
- Moberg, S.P. et al. (2019) Racial attitudes in America. *Public Opin. Q.* 83, 450–471
- Reardon, S.F. and Owens, A. (2014) 60 years after Brown: trends and consequences of school segregation. *Annu. Rev. Sociol.* 40, 199–218
- Anderson, K.M., ed (2012) *How Far Have We Come in Reducing Health Disparities? Progress Since 2000: Workshop Summary*, National Academies Press
- Gawronski, B., Payne, B.K., eds (2011) *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*, Guilford Press
- Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol. Rev.* 102, 4–27
- Axt, J.R. et al. (2014) The rules of implicit evaluation by race, religion, and age. *Psychol. Sci.* 25, 1804–1815
- Pasek, J. et al. (2009) Determinants of turnout and candidate choice in the 2008 US presidential election: illuminating the impact of racial prejudice and other considerations. *Public Opin. Q.* 73, 943–994
- Banaji, M.R. (2001) Implicit attitudes can be measured. In *The Nature of Remembering: Essays in Honor of Robert G. Crowder* (Roediger, H.L. III, et al., eds), pp. 117–150, American Psychological Association
- Devine, P.G. (1989) Stereotypes and prejudice: their automatic and controlled components. *J. Pers. Soc. Psychol.* 56, 5–18
- Bargh, J.A. (1999) The cognitive monster: the case against the controllability of automatic stereotype effects. In *Dual Process Theories in Social Psychology* (Chaiken, S. and Trope, Y., eds), pp. 361–382, Guilford Press

## Outstanding questions

What features in the environment cue implicit bias? Research has identified a range of environmental features associated with bias. The relative importance of various features at different levels of analysis, from workplaces to cities to nations, remains to be clarified.

Is implicit bias a summary of the past, or a prediction of the future? Theories that emphasize evaluative learning assume that implicit bias reflects that learning history. Predictive coding theories suggest that expectations about groups in the immediate future should drive implicit bias. All else being equal, past experience is a good predictor of future experience. However, experiments that dissociate past experience and predictions of the future can disentangle these possibilities.

When people change contexts, do their implicit biases change with them? Most of the research on real-world contexts has relied on correlational data, in which people inhabit different contexts. Laboratory experiments show that changing contexts can shift implicit bias. However, we know little about how implicit biases among the same population of people change as they shift from one context to another.

To what extent does implicit bias function as a marker of systemic inequalities, and to what extent does it function as a cause? These options are not mutually exclusive, as they may operate in self-reinforcing cycles. Future research should seek to examine these potential mutual influences.

Research suggests that implicit bias can be changed temporarily. Can shifting the temporary accessibility of implicit bias at strategic points, such as just before a decision is made, reduce biased decisions?



12. Rydell, R.J. *et al.* (2006) Of two minds: forming and changing valence-inconsistent implicit and explicit attitudes. *Psychol. Sci.* 17, 954–958
13. Lai, C.K. *et al.* (2016) Reducing implicit racial preferences: II. Intervention effectiveness across time. *J. Exp. Psychol. Gen.* 145, 1001–1016
14. Forscher, P.S. *et al.* (2019) A meta-analysis of procedures to change implicit measures. *J. Pers. Soc. Psychol.* 117, 522–559
15. Greenwald, A.G. (2015) Statistically small effects of the implicit association test can have societally large effects. *J. Pers. Soc. Psychol.* 108, 553–561
16. Devine, P.G. *et al.* (2012) Long-term reduction in implicit race bias: a prejudice habit-breaking intervention. *J. Exp. Soc. Psychol.* 48, 1267–1278
17. Greenwald, A.G. *et al.* (2009) Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *J. Pers. Soc. Psychol.* 97, 17–41
18. Oswald, F.L. *et al.* (2013) Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. *J. Pers. Soc. Psychol.* 105, 171–192
19. Cameron, C.D. *et al.* (2012) Sequential priming measures of implicit social cognition: a meta-analysis of associations with behavior and explicit attitudes. *Personal. Soc. Psychol. Rev.* 16, 330–350
20. Brownstein, M. *et al.* (2020) Understanding implicit bias: putting the criticism into perspective. *Pac. Philos. Q.* 101, 276–307
21. Gawronski, B. *et al.* (2017) Temporal stability of implicit and explicit measures: a longitudinal analysis. *Personal. Soc. Psychol. Bull.* 43, 300–312
22. Cooley, E. and Payne, B.K. (2017) Using groups to measure intergroup prejudice. *Personal. Soc. Psychol. Bull.* 43, 46–59
23. Cicchetti, D.V. (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* 6, 284
24. Payne, B.K. *et al.* (2017) The bias of crowds: how implicit bias bridges personal and systemic prejudice. *Psychol. Inq.* 28, 233–248
25. Nosek, B.A. *et al.* (2009) National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. U. S. A.* 106, 10593–10597
26. Hehman, E. *et al.* (2018) Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Soc. Psychol. Personal. Sci.* 9, 393–401
27. Leitner, J.B. *et al.* (2016) Racial bias is associated with ingroup death rate for Blacks and Whites: insights from Project Implicit. *Soc. Sci. Med.* 170, 220–227
28. Orchard, J. and Price, J. (2017) County-level racial prejudice and the black-white gap in infant health outcomes. *Soc. Sci. Med.* 181, 191–198
29. Rae, J.R. *et al.* (2015) Exposure to racial out-groups and implicit race bias in the United States. *Soc. Psychol. Personal. Sci.* 6, 535–543
30. Fazio, R.H. *et al.* (1995) Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *J. Pers. Soc. Psychol.* 69, 1013–1027
31. Fazio, R.H. *et al.* (1986) On the automatic activation of attitudes. *J. Pers. Soc. Psychol.* 50, 229–238
32. Higgins, E.T. (1996) Knowledge activation: accessibility, applicability, and salience. In *Social Psychology: Handbook of Basic Principles* (Higgins, E.T. and Kruglanski, A.W., eds), pp. 133–168, Guilford Press
33. Anderson, J.R. and Schooler, L.J. (1991) Reflections of the environment in memory. *Psychol. Sci.* 2, 396–408
34. Dixon, T.L. and Linz, D. (2000) Race and the misrepresentation of victimization on local television news. *Commun. Res.* 27, 547–573
35. Sonnett, J. *et al.* (2015) Priming implicit racism in television news: visual and verbal limitations on diversity. *Sociol. Forum* 30, 328–347
36. Payne, B.K. and Correll, J. (2020) Race, weapons, and the perception of threat. *Adv. Exp. Soc. Psychol.* 62, 1–50
37. Higgins, E.T. *et al.* (1982) Individual construct accessibility and subjective impressions and recall. *J. Pers. Soc. Psychol.* 43, 35–47
38. Higgins, E.T. *et al.* (1977) Category accessibility and impression formation. *J. Exp. Soc. Psychol.* 13, 141–154
39. Dasgupta, N. and Greenwald, A.G. (2001) On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *J. Pers. Soc. Psychol.* 81, 800–814
40. Lowery, B.S. *et al.* (2001) Social influence effects on automatic racial prejudice. *J. Pers. Soc. Psychol.* 81, 842–855
41. Brauer, M. *et al.* (2012) Describing a group in positive terms reduces prejudice less effectively than describing it in positive and negative terms. *J. Exp. Soc. Psychol.* 48, 757–761
42. Rudman, L.A. and Lee, M.R. (2002) Implicit and explicit consequences of exposure to violent and misogynous rap music. *Group Process. Intergroup Relat.* 5, 133–150
43. Lai, C.K. *et al.* (2013) Reducing implicit prejudice. *Soc. Personal. Psychol. Compass* 7, 315–330
44. Dasgupta, N. (2013) Implicit attitudes and beliefs adapt to situations: a decade of research on the malleability of implicit prejudice, stereotypes, and the self-concept. *Adv. Exp. Soc. Psychol.* 47, 233–279
45. Melnikoff, D.E. *et al.* (2020) Attitudes as prepared reflexes. *J. Exp. Soc. Psychol.* 88, 103950
46. Melnikoff, D.E. and Bailey, A.H. (2018) Preferences for moral vs. immoral traits in others are conditional. *Proc. Natl. Acad. Sci. U. S. A.* 115, E592–E600
47. Van Dessel, P. *et al.* (2019) Changing deep-rooted implicit evaluation in the blink of an eye: negative verbal information shifts automatic liking of Gandhi. *Soc. Psychol. Personal. Sci.* 10, 266–273
48. Cone, J. *et al.* (2017) Changing our implicit minds: how, when, and why implicit evaluations can be rapidly revised. *Adv. Exp. Soc. Psychol.* 56, 131–199
49. Fazio, R.H. and Olson, M.A. (2003) Implicit measures in social cognition research: their meaning and use. *Annu. Rev. Psychol.* 54, 297–327
50. Machery, E. (2017) Do indirect measures of biases measure traits or situations? *Psychol. Inq.* 28, 288–291
51. Schimmack, U. (2021) The Implicit Association Test: a method in search of a construct. *Perspect. Psychol. Sci.* 16, 396–414
52. Galton, F. (1907) Vox populi. *Nature* 75, 450–451
53. Mellers, B. *et al.* (2014) Psychological strategies for winning a geopolitical forecasting tournament. *Psychol. Sci.* 25, 1106–1115
54. Hehman, E. *et al.* (2019) Establishing construct validity evidence for regional measures of explicit and implicit racial bias. *J. Exp. Psychol. Gen.* 148, 1022–1040
55. Murphy, M.C. and Walton, G.M. (2013) From prejudiced people to prejudiced places: a social-contextual approach to prejudice. In *Frontiers of Social Psychology. Stereotyping and Prejudice* (Stangor, C. and Crandall, C.S., eds), pp. 181–203, Psychology Press
56. Murphy, M.C. *et al.* (2018) Prejudiced places: how contexts shape inequality and how policy can change them. *Policy Insights Behav. Brain Sci.* 5, 66–74
57. Farkas, I. *et al.* (2002) Mexican waves in an excitable medium. *Nature* 419, 131–132
58. Gawronski, B. and Bodenhausen, G.V. (2017) Beyond persons and situations: an interactionist approach to understanding implicit bias. *Psychol. Inq.* 28, 268–272
59. Rivers, A.M. *et al.* (2017) Implicit bias reflects the personal and the social. *Psychol. Inq.* 28, 301–305
60. Connor, P. and Evers, E.R. (2020) The bias of individuals (in crowds): why implicit bias is probably a noisily measured individual-level construct. *Perspect. Psychol. Sci.* 15, 1329–1345
61. Payne, B.K. *et al.* Reply to Connor and Evers (2020): critique of the bias of crowds model simply restates the model. *Perspect. Psychol. Sci.* Published online December 29, 2020. <http://doi.org/10.31234/osf.io/ph7yt>.
62. Rudman, L.A. *et al.* (2007) Developmental sources of implicit attitudes. *Personal. Soc. Psychol. Bull.* 33, 1700–1713
63. Vuletic, H.A. and Payne, B.K. (2019) Stability and change in implicit bias. *Psychol. Sci.* 30, 854–862
64. Payne, B.K. and Vuletic, H.A. (2018) Policy insights from advances in implicit bias research. *Policy Insights Behav. Brain Sci.* 5, 49–56

65. Crandall, C.S. *et al.* (2018) Changing norms following the 2016 US presidential election: the Trump effect on prejudice. *Soc. Psychol. Personal. Sci.* 9, 186–192
66. Tankard, M.E. and Paluck, E.L. (2016) Norm perception as a vehicle for social change. *Soc. Issues Policy Rev.* 10, 181–211
67. Payne, B.K. *et al.* (2019) Historical roots of implicit bias in slavery. *Proc. Natl. Acad. Sci. U. S. A.* 116, 11693–11698
68. Charlesworth, T.E. and Banaji, M.R. (2019) Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016. *Psychol. Sci.* 30, 174–192
69. Sadler, M. and Devos, T. (2020) Ethnic diversity matters: putting implicit associations between weapons and ethnicity in context. *Group Process. Intergroup Relat.* 23, 285–300
70. Somo, A. *et al.* (2020) Implicit black-weapon associations weakened over time in increasingly multiethnic metropolitan areas. *Anal. Soc. Issues Public Policy* Published online December 18, 2020. <https://doi.org/10.1111/asap.12228>
71. Devos, T. *et al.* (2019) Temporal fluctuations in context ethnic diversity over three decades predict implicit national inclusion of Asian Americans. *Group Process. Intergroup Relat.* Published online November 25, 2019. <https://doi.org/10.1177/1368430219887440>
72. Ofosu, E.K. *et al.* (2019) Same-sex marriage legalization associated with reduced implicit and explicit antigay bias. *Proc. Natl. Acad. Sci. U. S. A.* 116, 8846–8851
73. Ravary, A. *et al.* (2019) Shaping the body politic: mass media fat-shaming affects implicit anti-fat attitudes. *Personal. Soc. Psychol. Bull.* 45, 1580–1589