

# The Blind Leads the Blind: Denying Systemic Discrimination in Discussions of Implicit Bias

*Aspen Ho, Simon Fraser University*

## **Abstract**

This paper was originally written for Dr. Alyssa Croft's PSYC 300W course *Critical Analysis of Issues in Psychology*. The assignment asked students to critically analyze a scholarly trade book about the psychology of success. In the style of a book report, the term paper explored how well the trade book represents the research it cites. Students then compared two new empirical articles – one supporting and one refuting the book's argument – before explaining which article was more convincing. The assignment concludes with the student's overall opinion of the book. The paper uses APA citation style.

Banaji and Greenwald (2013) claim that implicit biases – automatic, prejudiced associations – are consciously inaccessible and responsible for modern discrimination. However, the supposed inaccessibility of implicit bias removes blame for resulting discriminatory actions. The authors argue that implicit bias results from simple pattern recognition – denying centuries of deliberate, systemic discrimination. In this paper, I compare the methods of two articles on implicit bias awareness. The first article portrays implicit bias as unconscious, using explicit bias measures about consciously held beliefs. The article posits that lower explicit bias scores imply low awareness of implicit bias. However, conscious disapproval of implicit bias does not suggest obliviousness. Conversely, the opposing article directly measures awareness by asking participants to predict their implicit bias scores. Highly accurate predictions presented implicit bias as accessible to conscious awareness. The opposing article's compelling findings and robust methods refute Banaji and Greenwald's (2013) unconscious claim, reinstating responsibility onto oppressors for implicit discrimination.

## **Book Summary**

Banaji and Greenwald (2013) claim that most modern discrimination results from implicit biases of unaware, righteous people. Implicit biases are prejudiced associations inaccessible to conscious awareness, often contradicting more explicit egalitarian beliefs. Success entails reducing implicit discriminatory behaviour with policies that mitigate stereotype activation (Banaji & Greenwald, 2013). For example, blind hiring processes can block discrimination by obscuring identifying features of marginalized identities.

The authors use studies with the Implicit Association Test (IAT) to demonstrate how well-meaning individuals can have unconscious preferences for the dominant social class. The classic race IAT measures reaction times of associating White or Black people with positive or negative adjectives. Participants quicker to associate White people with positive adjectives and Black people with negative adjectives show implicit White preference (Banaji & Greenwald, 2013). A prejudiced IAT score predicts subtle discriminatory behaviours in healthcare, employment, housing, policing, and other areas. The authors claim that these behaviours do not constitute racism because they are not purposeful.

Banaji and Greenwald (2013) affably combine research, personal anecdotes, and analogies to illustrate the behavioural effects of implicit bias. They cater to an egalitarian audience largely unfamiliar with subtle forms of prejudice. The authors extend a permissive, welcoming hand to individuals feeling defensive about carrying implicit biases.

### **Critical Analysis**

Portraying implicit bias as completely outside a person's awareness and control removes moral responsibility for discriminatory behaviour. Banaji and Greenwald (2013) reduce racial prejudice to an unavoidable product of pattern recognition, ignoring the intentional and historical origins of systemic discrimination. Because they view implicit bias as consciously inaccessible, Banaji and Greenwald (2013) express their support for blinding methods over long-term anti-racism training. They believe that individuals cannot change their discriminatory behaviour, thus they encourage policies that instead hide targets of discrimination. This permits racist actions from oppressors who supposedly can't help themselves. This puts responsibility onto marginalized communities to assimilate or risk harm. While claiming to uncover implicit bias, Banaji and Greenwald (2013) instead enforce the prejudiced system.

The authors use their claim of unconscious bias to perpetuate a narrow, removed view of racism. For example, they dilute the concept of the "aversive

racist” – one who avoids interacting with people of other races – to the “uncomfortable egalitarian”, refusing to view racism as anything less than purposeful violence or extremist hate speech. In the appendix, the authors finally mention unconscious discrimination as possibly, partially causing Black disadvantage. Banaji and Greenwald (2013) attribute their silence to the apparent impossibility of getting decisive evidence on this complex issue. Their empirical stance is flagrantly hypocritical, for Banaji and Greenwald (2013) often misconstrue evidence in their own book. They claim that participants in McConahay et al. (1981) unconsciously changed their questionnaire responses between White and Black experimenters, implicitly expressing fewer racist beliefs around a Black person. However, McConahay et al. (1981) clearly state that participants consciously, intentionally altered their responses according to the experimenter’s race. Banaji and Greenwald’s (2013) description of this study is blatantly false, framing implicit bias as inaccessible and unworkable. The authors’ hypocrisy and overall ignorance support a superficial, removed misrepresentation of human rights issues.

#### **Fact Check A: Cunningham et al. (2004)**

##### ***Article Summary***

Cunningham et al. (2004) differentiate between implicit and explicit types of prejudice and ethnocentrism, operationalizing these attitudes via various questionnaires and the IAT. The researchers define ethnocentrism as an overall preference for one’s own social group. Cunningham et al. (2004) found that unconscious, implicit prejudice and conscious, explicit prejudice were different yet related by ethnocentrism. This is the first study to find latent correlation between implicit and explicit prejudice, allowing both constructs to interact within a general ethnocentric attitude. However, Cunningham et al. (2004) argue that implicit and explicit prejudice are still distinct because the IAT showed larger bias than self-report measures, implying that only explicit attitudes are consciously accessible.

##### ***Critical Analysis***

Cunningham et al. (2004) did well in exposing hidden relationships between implicit and explicit bias, but they fail to support the construct validity of their measures. They claim that implicit bias is completely inaccessible to awareness because the IAT reports greater bias than explicit measures, but they never explore the perceptual experience of taking the IAT. Explicit self-report scales

measure conscious, thought-out beliefs, but not whether participants notice their contradicting automatic associations. In other words, “automatic” is not necessarily synonymous with “inaccessible”. Participants may perceive rapid bias activation independent of whether they endorse those beliefs or know how to override them.

This study convincingly connects group-specific prejudice to general ethnocentrism. Historically, the IAT has compared preferences for two groups – such as Black and White people – but Cunningham et al. (2004) use a broader approach by measuring ethnocentrism. Their findings demonstrate that prejudice against one group does not exist in a vacuum; rather, high levels of bias in one regard likely reflect an overarching preference for one’s own social class. In short, the researcher’s methods were generally appropriate but could have benefitted from verifying the inaccessibility of implicit bias.

### ***Fact Check***

Cunningham et al. (2004) also explore how rigid attitudes relate to ethnocentrism and right-wing ideology, but these variables were irrelevant for my purposes. They found that stronger supporters of right-wing ideology likely had highly ethnocentric attitudes and simpler, more rigid cognitive styles. I omitted these findings in favour of expanding on their results about implicit and explicit prejudice. The contents of Cunningham et al. (2004) were thus greatly reduced in my summary to prevent dilution of their relevant claims.

I also omitted many methodological details and analytical logistics for brevity’s sake. Cunningham et al. (2004) corrected for the IAT’s measurement error via latent variable analysis, finding correlations between the IAT and explicit measures that were invisible to other studies. To avoid confusing clutter, my summary did not specify how Cunningham et al. (2004) yielded different results than previous data.

### **Fact Check B: Hahn et al. (2014)**

#### ***Article Summary***

Contrary to claims that implicit bias is consciously inaccessible, Hahn et al. (2014) hypothesize that individuals can perceive and evaluate their own implicit bias. Before taking the IAT, participants predicted their bias scores with significant accuracy. Their success implies reliable access to implicit bias. Prediction accuracy persisted regardless of whether participants believed that IAT scores merely reflected their cultural surroundings versus revealed their true feelings;

participants could report their implicit bias despite the threat to their self-presentation. Hahn et al. (2014) also replicated low correlations from previous studies between the IAT and explicit measures. They explain that while implicit biases exist within awareness, explicit measures allow participants to select which associations accurately represent their conscious beliefs. Explicit measures thus do not measure the awareness of implicit bias.

### ***Critical Analysis***

Hahn et al. (2014) effectively refute previous claims of implicit bias's inaccessibility by overcoming the complications of divergent measures. Mismatched scores between implicit and explicit measures do not sufficiently imply unawareness of implicit bias, but previous studies had ignored this issue. Hahn et al. (2014) directly measure participants' awareness of their implicit bias to counteract assumptions drawn from dissociations between methods.

There is little convergent validity between different measures of implicit bias, thus Hahn et al.'s (2014) findings are limited to the IAT. They justify their narrow scope with the IAT's widespread use in literature. However, the mysteriously low convergent validity suggests that different measures target different aspects of implicit bias. Because Hahn et al. (2014) focus on implicit bias overall, using other implicit measures along with the IAT would improve the operationalization of their target construct. Overall, the experimenters successfully investigate awareness of implicit bias as measured by the IAT, but their study would benefit from including a greater variety of implicit measures.

### ***Fact Check***

Hahn et al. (2014) tested IAT prediction accuracy across several conditions, varying the type of prediction scales to the amount of IAT training participants received prior. I omitted these details because predictions remained accurate across all conditions. Additionally, I did not explain that Hahn et al. (2014) tested several prediction scales during the experiment, altering their methods for possible confounding variables. This trial-and-error information is irrelevant for my purposes.

My summary effectively connects the study's findings to a larger social context. In my Critical Analysis of Banaji and Greenwald (2013), I establish the harms of claiming that implicit bias is consciously inaccessible. Emphasizing how Hahn et al. (2014) discredits this claim thus supports my own argument while maintaining accuracy to the original study.

### **Article Comparison**

Hahn et al. (2014) more convincingly support their argument that implicit bias is within awareness. Cunningham et al. (2004) rely only on explicit measures to detect awareness of both explicit and implicit bias, whereas Hahn et al. (2014) show that explicit measures largely miss implicit bias awareness. Hahn et al. (2014) demonstrate a more specific, critical approach to construct validity. They make minimal assumptions about the properties of implicit bias, instead directly testing awareness across various conditions. The researchers clearly outline how certain conditions may remove confounding variables or reveal different aspects about implicit bias.

Cunningham et al. (2004) attempt a broader approach, testing for prejudice, ethnocentrism, rigidity, and right-wing ideology. Unfortunately, they lack the meticulous research needed for their claims. For example, they differentiate between implicit and explicit bias using the low correlation of their measures, failing to provide evidence for the separate properties of bias types. They instead merely assume that the rest of their theoretical model is true, asserting explicit bias as conscious and implicit bias as unconscious (Cunningham et al., 2004). Ultimately, Hahn et al. (2014) present a more compelling argument with a narrow yet thorough exploration of implicit bias awareness.

### **Overall Book Impressions**

I chose this book because I am very passionate about subtle forms of discrimination and the systemic perpetuation of disadvantage. I believe that working through internal prejudice is difficult but rewarding, especially when privilege may blind someone to the inequalities that others suffer. For these reasons, I absolutely despised this book. I had hoped to learn more actionable ways to combat unseen discrimination, but I instead encountered everything from denials of systemic racism to homophobic jokes. Regarding their personal solutions for reducing discrimination, Banaji and Greenwald (2013) describe a single action: they change their screensaver to show people of different races. This apparently helps remind them of the world's diversity. Their anecdote exposes a reductive, sheltered view of racial injustice. Perhaps going outside their lab to engage with disadvantaged communities might better remind them of diversity. I would not recommend this book to my worst enemy, but that's impossible because my worst enemies wrote this book.

## References

Banaji, M.R. & Greenwald, A.G. (2013). *Blindspot: Hidden Biases of Good People*. Delacorte Press.

Cunningham, W.A., Nezlek, J.B., & Banaji M.R. (2004). Implicit and explicit ethnocentrism: Revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin*, 30(10), 1332-1346.  
<https://doi.org/10.1177/0146167204264654>

Hahn, A., Judd, C.M., Hirsh, H.K., & Blair, I.V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, 143(3), 1369–1392.  
<https://doi.org/10.1037/a0035028>

## AI Transparency Statement

No AI was used to complete this assignment.

By submitting this essay, I attest that it is my own work, completed in accordance with University regulations. I also give permission for the Student Learning Commons to publish all or part of my essay as an example of good writing in a particular course or discipline, or to provide models of specific writing techniques for use in teaching. This permission applies whether or not I win a prize, and includes publication on the Simon Fraser University website or in the SLC Writing Contest Open Journal.

This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

© Aspen Ho, 2024

Available from: <https://journals.lib.sfu.ca/index.php/slc-uwcc>