

Implicit Bias Is a Public-Health Problem, and Hearts and Minds Are Part of the Solution

Michael A. Olson and Laura J. Gill 

Department of Psychology, University of Tennessee

Greenwald, Dasgupta, Dovidio, Kang, Moss-Racusin, and Teachman (this issue) distinguish among individual-, group-, and institutional-level interventions to reduce discrimination. We refer to the first two as “hearts and minds” approaches because they are designed to mitigate or reduce prejudice within individuals. Institutional-level interventions, on the other hand, attempt to render individual biases irrelevant to decision processes through organizational policies and practices. On the basis of the available evidence, the authors of the target article doubt the efficacy of hearts-and-minds approaches and emphasize the prospects of institutional-level approaches from a preventative public-health perspective. As we are all steeped in the same literatures, we sympathize with the authors’ doubt. However, and despite both social and scientific attention away from individual bias and toward systemic bias, we argue that it is difficult to create systemic change by bypassing hearts and minds and that interventions attempting to change them should be preserved and refined. To that end, we offer a framework that we hope will advance research and interventions that target groups and individuals. Ultimately, we encourage an integration of individual-level and public-health perspectives to address implicit bias.

Before presenting our case, it is worth noting that we have few substantive disagreements with the authors’ take on the literature as it presently sits. We are in general agreement that implicit bias is pervasive and influential, that the Implicit Association Test (IAT) is a reasonably good tool with which assess it, and that implicit bias and explicit bias are not necessarily distinct entities in people’s heads (cf. Wilson et al., 2000). We also agree, and echo work by Paluck (e.g., Paluck et al., 2021) and Lai (e.g., Lai et al., 2016), that interventions designed to weaken implicit bias (or prejudice generally) tend to disappoint once they leave the confines of a single laboratory session. Our points of departure from the authors include how implicit bias ought to be conceptualized and its modes of influence on judgments and behavior. Each of these have implications for interventions aimed at both mitigating bias’s impact and

reducing bias itself. We also offer insight into two questions posed by the authors for which they believe there is insufficient data to answer: (a) “Are implicit biases introspectively (consciously) accessible?” (we think they are) and (b) “How do the association strengths measured by the IAT influence social behavior?” (we will explain below). The dual-process framework we describe may help resolve some of these disagreements, illuminate when and how individual implicit bias leads to discrimination, and explain how interventions might be improved to mitigate or reduce it.

A primary point we hope to convey is that the institutional policies and practices informed by public-health perspectives are not immune to the effects of biased individual hearts and minds. Policies and practices are proposed, adopted, enforced, and evaluated by individuals. Without buy-in by individuals who have power within organizations, policies and practices are likely to be inconsistently enforced or actively undermined (Dobbin & Kalev, 2016; Pearce & Conger, 2003). People wish to see value in their work and do not like to be told what to do. Interventions that do not seek buy-in, emphasize what not to do, or are perceived to reduce individual autonomy are likely to backfire (e.g., Does et al., 2011; Legault et al., 2011). In fact, from a historical perspective, some argue that institutional change has been driven more by individuals at the grass-roots level than by institutional leadership (e.g., Scott, 2012). For example, social-movement theory recognizes the limitations of top-down-only approaches and highlight the role of “tempered radicals,” who work within organizations to effect change from the bottom up, even without formal power (Meyerson, 2003).

Clarifying “Implicit”

What is “implicit bias”? Our view is that the authors’ two criteria are both problematic. The first refers solely

Corresponding Author:

Michael Olson, Department of Psychology, University of Tennessee
Email: olson@utk.edu

to measurement: it “manifests (and can be measured) *indirectly* [emphasis added]” (Greenwald et al., 2022, p. 8). The second refers to conscious access: “[it] can operate without those who perpetrate discrimination needing to be aware either of their biased associations or of the role those associations play in guiding their judgment and action” (Greenwald et al., 2022, p. 8). Regarding the first criteria, much has been written about the problem of conflating measurement outputs with psychological constructs (De Houwer et al., 2009; Fazio & Olson, 2003). Should an attitude be considered “implicit” merely because it was assessed implicitly? If asked directly (i.e., explicitly), we will say we like chocolate. If an IAT also reveals in us a prochocolate bias, what have we learned that is new? It is more parsimonious to simply say that we like chocolate. That is, outputs of implicit measures should not be used to infer how an attitude is represented in memory (or, we would add, how it was acquired) or whether it is distinct from some other kind(s) of attitudes.

Regarding the second criteria (and one of the authors’ lingering questions), there is no evidence that people are unaware of attitudes assessed by implicit measures. Instead, there is evidence that people are aware of them. For example, we (Phillips & Olson, 2014) implored some participants to be as honest as possible while completing explicit measures of racial prejudice after completing an IAT measure of race bias. Other participants simply completed the IAT and then the explicit measures. In the former condition, IAT scores and explicit measures were correlated; in the latter condition, they were not. Similar results have been reported in both prejudice-related domains and other domains (e.g., Hahn & Gawronski, 2019; Olson et al., 2007). Thus, it appears people have conscious access to their implicit biases. So, what distinguishes “implicit bias” from explicit bias?

The term “implicit” may have become too imprecise for productive science; some authors have recommended abandoning the term entirely (Corneille & Hütter, 2020). Unfortunately, it is probably too enmeshed in popular discourse to renounce. Arguably, its most useful connotation regards the automaticity of the attitude—how quickly, efficiently, and inescapably it is activated upon perception of the target (Fazio, 2007). Conveniently for these debates, implicit measures such as the IAT and evaluative priming effectively gauge the extent to which an attitude has these automatic properties (Olson & Zabel, 2016). For the present purposes, if we equate “implicit attitude/bias” with “automatic attitude/bias,” we can agree that the IAT measures implicit bias. To be clear, the critical point is that to be a useful designation, *implicit* should refer to qualities of the attitude, not whether it was measured with an

IAT or whether people are aware of the attitude itself. We will use the term *implicit bias* to refer to prejudices with automatic properties, and, on the basis of our reading of their work, we suspect the authors of the target article would generally agree that the IAT reveals these properties.

A Dual-Process Approach

Within our dual-process framework (motivation and opportunity as determinants [MODE] model; Fazio, 1990; Fazio & Olson, 2014), implicit bias is the starting point of the attitude-judgment/behavior relation. If the bias is sufficiently strong, mere perception of a relevant target (e.g., a member of a minoritized group) will activate the bias from memory spontaneously and inescapably. Once activated, implicit bias can guide attention toward the target, bias construal of it, and prompt approach/avoidance tendencies toward or away from it (Fazio, 2007). Much social life probably proceeds this way, relatively thoughtlessly. Critically, although people are largely aware of their implicit biases, they may be unaware of their *impact* on basic perception (Gawronski et al., 2006). Raising awareness of the impact of bias is one way in which hearts-and-minds approaches, particularly diversity training (DT), can contribute to reductions in discriminatory behavior (e.g., Carnes et al., 2015).

The MODE model also posits motivation and opportunity as determinants of whether implicit bias guides behavior spontaneously or its impact is moderated. Motivation can be broad (e.g., accuracy, affiliation) or specific (motivation to avoid prejudice, motivation to avoid disapproval; Dunton & Fazio, 1997; Plant & Devine, 1998). Often, relevant motives are at odds with implicit bias; for example, aversive racism is characterized by implicit bias but also by motivation to avoid appearing biased to oneself or others (Pearson et al., 2009). Given the pervasiveness of both automatic prejudice and such good intentions, aversive forms of prejudice are probably widespread.

Critically, motivation is able to overcome the impact of implicit bias only with sufficient opportunity. Preventing bias’s influence takes effort, and opportunity includes any variable that affects one’s ability to exert that effort. For example, time constraints, fatigue, distraction, cognitive load, and intoxication all impede deliberative, motivated processing. Thus, for motivation to succeed in overcoming bias, opportunity is essential; without sufficient opportunity, judgments and behavior will be bias-driven regardless of motivation.

Our goal here was merely to provide a brief overview of the model (for a more detailed treatment and supporting evidence, see Fazio & Olson, 2014; Olson

& Fazio, 2008). For the present purposes, we believe the MODE model clarifies a number of debates surrounding implicit bias, including the authors' first lingering question, "How do the association strengths measured by the IAT influence social behavior?" As we briefly reviewed, it does so by clarifying the processes involved and the conditions under which bias affects judgments and behaviors.

Another debate regards the unfortunately vague question of just how much implicit bias matters in terms of discriminatory behavior. As the authors review, a number of meta-analyses reveal implicit bias-behavior correlations ranging from small to moderate in size. We agree that even small effects can have large impact for the reasons the authors describe. That said, as attitude-behavior researchers, we are frustrated that the debate centers on "first generation" questions of *whether* a bias-behavior relationship exists when the field is advanced enough to ask more informative questions regarding *when* implicit bias affects behavior: We really like chocolate, but we do not always eat it whenever it is available. That attitudes should *always* predict behavior is untenable: We might be motivated to reduce our sugar intake or to avoid the ire of our children after we have deprived them of chocolate. The MODE model makes clear predictions about when implicit bias should affect judgments and behavior: Specifically, when either motivation or opportunity is low. In such contexts, implicit bias reliably predicts judgments and behavior (Fazio & Olson, 2014). Thus, our position is that implicit bias can matter quite a lot, but it depends on motivation and opportunity.

The MODE model also provides insight into debates surrounding relationships between implicit and explicit measures. A response on a survey is an explicit, deliberate behavior—one chooses which box to check or not check. As relatively "downstream" behaviors, responses on surveys might be influenced by implicit bias, motivation, or some combination of both. Consistent with the MODE model, when a respondent is not motivated to avoid bias, or lacks the opportunity to correct for bias (e.g., their cognitive resources are compromised; Loersch et al., 2015), implicit measures and explicit measures tend to align. When one is motivated to correct for bias and afforded the opportunity to do so, implicit and explicit measures tend to diverge. Thus, for the MODE model, implicit measures get at something closer to "the truth," or at least one's gut feelings (Dijksterhuis et al., 2008), whereas explicit measures are variably influenced by both gut feelings and motives. A further implication of the MODE model is that people do not have two attitudes, one implicit and one explicit (cf. Wilson et al., 2000). Instead, they have gut feelings (implicit biases) and motives (derived

from beliefs) that may sometimes be in conflict, as in the case of aversive racism.

Implications for Interventions

We agree that the individual-level prejudice reduction literature is disappointing, but the group-level (i.e., DT) literature is even more so. The former offers only limited insight into how prejudice in general and implicit bias in particular might be "durably modified," but most of this research is at least systematic and theory-driven. The same cannot be said of the DT literature, where there is consensus neither on what such training should entail nor what its expected outcomes should be (apart from the ultimate goal of increasing representation or retention of minoritized groups); outcomes of DT research are laundry lists with little theoretical organization (Anand & Winters, 2008; Atewologun et al., 2018). We believe the MODE model can provide some.

The MODE model identifies three key foci for interventions: (a) increase awareness, not of bias itself (because people are generally aware), but of the impact of bias; (b) increase motivation to counteract the effects of bias; and (c) increase opportunity for motivated processes to override implicit ones. Critically, all three are necessary to reduce the impact of bias.

Increasing awareness of the impact of bias is largely an information-delivery endeavor, and several studies provide evidence that DT and other interventions are effective at achieving it (e.g., Carnes et al., 2015; Moss-Racusin et al., 2016). Again, awareness of bias's impact is necessary to prevent its influence, but it is not sufficient; one needs motivation and opportunity as well. Thus, it is not surprising that DT often shows effects on awareness but not behavior.

Increasing motivation can entail two distinct approaches based on two well-documented prejudice-related motives: internal motivation, which is based on personally important egalitarian standards, and external motivation, which is based on external demands and rewards (Plant & Devine, 1998). Internal motivation is increased through "good contact," the development of personal relationships and acceptance from out-group members, and the internalization of out-group goals (Kunstman et al., 2013; Tropp & Barlow, 2018). External motivation is driven largely by normative concerns, and classic social psychological research indicates that people generally comply with strong social norms. Normative interventions have proven effective in laboratory and field studies at preventing the expression of prejudice (e.g., Murrar et al., 2020). No research to our knowledge has examined the impact of DT on perceptions of antiprejudice norms or their impact on discrimination, and few DT ventures employ

evidence-based approaches toward a direct goal of increasing motivation per se.

Opportunity can have psychological bases (e.g., distraction, fatigue) or structural ones (e.g., work-related stressors, time pressure), also implying distinct interventions. As of this writing, the COVID-19 pandemic persists; climate change is accelerating; Russia has invaded Ukraine; anxiety, depression, suicide, and substance use are all rising to record levels; and bias is still rampant. Work-related stressors have increased, and those charged with hiring decisions are provided less time and support than they would like to make fair and effective ones. All of these constraints affect the opportunity to think deliberatively about bias and avoid its impact. It is easier said than done, of course, but well-rested people are less rash, and accessible care for physical and mental health, adequate pay, and better work-life balance would afford greater opportunity for motivated processes to counteract bias. Thus, we welcome structural interventions that increase individual-level opportunity.

And what of attempts to durably modify implicit bias itself? Here we echo points made in the target article, particularly regarding longitudinal approaches. For example, in addition to Teachman and Woody's (2003) longitudinal work effecting change in phobia-related associations, we mention our own longitudinal work using evaluative conditioning to durably modify implicit attitudes toward marriage partners (McNulty et al., 2017). So far, it appears that durable change in the information environment appears necessary for durable change in implicit bias.

Of course, one may also pose the question of whether public-health-inspired institutional interventions are capable of producing durable change. As we have argued, our sense is that without individual-level buy-in, it will be difficult to effectively implement institution-level interventions in a way that produces durable reductions in discrimination. For example, chief diversity officers' power to influence hiring and retention decisions will probably be limited by the extent to which their goals, decisions, and the position itself are valued by individuals within organizations.

Conclusion

Hearts-and-minds interventions should move beyond the one-size-fits-all approach of shepherding everyone to the same DT workshop. Instead, interventions should be targeted and follow a process analogous to public-health approaches that identify specific disparities and their causes in specific domains before implementing solutions. Such tailored approaches would acknowledge that some contexts and individuals lack strong

antiprejudice norms, some lack internal motivation, some lack awareness of the impact of bias, and some deprive people opportunity to make conscientious decisions. The MODE model provides guidance about how to identify these specific problems, and basic laboratory research provides guidance about how interventions might be tailored and applied to address them.

Finally, we would do well to observe the mutual influence exchanged between individuals and institutions and ought not consider one in isolation of the other (Pettigrew, 2021). Institutional policies and practices can affect individual behavior (especially when there is buy-in from individuals), and individuals can effect change in institutions (especially when they act collectively; Kezar, 2012). Indeed, meta-analyses suggest that such two-pronged approaches may be particularly effective at increasing representation of minoritized groups (e.g., Bezrukova et al., 2016). Group-level interventions may be uniquely suited to effect institutional change because they involve shared experiences between individuals, which can cultivate shared language, norms, and goals and facilitate collective action toward institutional change (Shteynberg et al., 2020). Ultimately, bottom-up and top-down approaches should meet in the middle: Institutional interventions should be coupled with hearts-and-minds interventions.

Transparency

Editor: Nora S. Newcombe

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

ORCID iD

Laura J. Gill  <https://orcid.org/0000-0002-4324-2806>

References

- Anand, R., & Winters, M. F. (2008). A retrospective view of corporate diversity training from 1964 to the present. *Academy of Management Learning & Education*, 7(3), 356–372. <https://doi.org/10.5465/amle.2008.34251673>
- Atewologun, D., Cornish, T., & Tresh, F. (2018). *Unconscious bias training: An assessment of the evidence for effectiveness* (Research Report 113). U.K. Equality and Human Rights Commission. <https://www.equalityhumanrights.com/sites/default/files/research-report-113-unconscious-bias-training-an-assessment-of-the-evidence-for-effectiveness-pdf.pdf>
- Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin*, 142(11), 1227–1274. <https://doi.org/10.1037/bul0000067>

- Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., Forscher, P., Isaac, C., Kaatz, A., Magua, W., Palta, M., & Sheridan, J. (2015). Effect of an intervention to break the gender bias habit for faculty at one institution: A cluster randomized, controlled trial. *Academic Medicine: Journal of the Association of American Medical Colleges*, 90(2), 221–230. <https://doi.org/10.1097/ACM.0000000000000552>
- Corneille, O., & Hütter, M. (2020). Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. *Personality and Social Psychology Review*, 24(3), 212–232. <https://doi.org/10.1177/1088868320911325>
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, 135(3), 347–368.
- Dijksterhuis, A., Albers, L. W., & Bongers, K. C. (2008). Digging for the real attitude: Lessons from research on implicit and explicit self-esteem. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 249–270). Psychology Press.
- Dobbin, F., & Kalev, A. (2016). Why diversity programs fail. *Harvard Business Review*, 94(7), Article 14. <https://hbr.org/2016/07/why-diversity-programs-fail>
- Does, S., Derks, B., & Ellemers, N. (2011). Thou shalt not discriminate: How emphasizing moral ideals rather than obligations increases Whites' support for social equality. *Journal of Experimental Social Psychology*, 47(3), 562–571. <https://doi.org/10.1016/j.jesp.2010.12.024>
- Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, 23(3), 316–326. <https://doi.org/10.1177/0146167297233009>
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60318-4](https://doi.org/10.1016/S0065-2601(08)60318-4)
- Fazio, R. H. (2007). Attitudes as object–evaluation associations of varying strength. *Social Cognition*, 25(5), 603–637. <https://doi.org/10.1521/soco.2007.25.5.603>
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54(1), 297–327. <https://doi.org/10.1146/annurev.psych.54.101601.145225>
- Fazio, R. H., & Olson, M. A. (2014). The MODE model: Attitude-behavior processes as a function of motivation and opportunity. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual process theories of the social mind* (pp. 155–171). Guilford Press.
- Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are “implicit” attitudes unconscious? *Consciousness and Cognition*, 15(3), 485–499. <https://doi.org/10.1016/j.concog.2005.11.007>
- Greenwald, A. G., Dasgupta, N., Dovidio, J. F., Kang, J., Moss-Racusin, C. A., & Teachman, B. A. (2022). Implicit-bias remedies: Treating discriminatory bias as a public-health problem. *Psychological Science in the Public Interest*, 23(1), 7–40. <https://doi.org/10.1177/15291006211070781>
- Hahn, A., & Gawronski, B. (2019). Facing one's implicit biases: From awareness to acknowledgment. *Journal of Personality and Social Psychology*, 116(5), 769–794. <https://doi.org/10.1037/pspi0000155>
- Kezar, A. (2012). Bottom-up/top-down leadership: Contradiction or hidden phenomenon. *The Journal of Higher Education*, 83(5), 725–760. <https://doi.org/10.1353/jhe.2012.0030>
- Kunstman, J. W., Plant, E. A., Zielaskowski, K., & LaCrosse, J. (2013). Feeling in with the outgroup: Outgroup acceptance and the internalization of the motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 105(3), 443–457. <https://doi.org/10.1037/a0033082>
- Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchard, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8), 1001–1016. <https://doi.org/10.1037/xge0000179>
- Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, 22(12), 1472–1477. <https://doi.org/10.1177/0956797611427918>
- Loersch, C., Bartholow, B. D., Manning, M., Calanchini, J., & Sherman, J. W. (2015). Intoxicated prejudice: The impact of alcohol consumption on implicitly and explicitly measured racial attitudes. *Group Processes & Intergroup Relations*, 18(2), 256–268. <https://doi.org/10.1177/1368430214561693>
- McNulty, J. K., Olson, M. A., Jones, R. E., & Acosta, L. M. (2017). Automatic associations between one's partner and one's affect as the proximal mechanism of change in relationship satisfaction: Evidence from evaluative conditioning. *Psychological Science*, 28(8), 1031–1040. <https://doi.org/10.1177/0956797617702014>
- Meyerson, D. E. (2003). *Tempered radicals: How everyday leaders inspire change at work*. Harvard Business School Press.
- Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2016). A “scientific diversity” intervention to reduce gender bias in a sample of life scientists. *CBE—Life Sciences Education*, 15(3), Article ar29. <https://doi.org/10.1187/cbe.15-09-0187>
- Murrar, S., Campbell, M. R., & Brauer, M. (2020). Exposure to peers' pro-diversity attitudes increases inclusion and reduces the achievement gap. *Nature Human Behaviour*, 4(9), 889–897. <https://doi.org/10.1038/s41562-020-0899-5>
- Olson, M. A., & Fazio, R. H. (2008). Implicit and explicit measures of attitudes: The perspective of the MODE model. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 19–64). Psychology Press.
- Olson, M. A., Fazio, R. H., & Hermann, A. D. (2007). Reporting tendencies underlie discrepancies between implicit and explicit measures of self-esteem. *Psychological*

- Science*, 18(4), 287–291. <https://doi.org/10.1111/j.1467-9280.2007.01890.x>
- Olson, M. A., & Zabel, K. L. (2016). Measures of prejudice. In T. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 175–212). Psychology Press.
- Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, 72, 533–560. <https://doi.org/10.1146/annurev-psych-071620-030619>
- Pearce, C., & Conger, J. (2003). *Shared leadership*. Sage.
- Pearson, A. R., Dovidio, J. F., & Gaertner, S. L. (2009). The nature of contemporary prejudice: Insights from aversive racism. *Social and Personality Psychology Compass*, 3(3), 314–338. <https://doi.org/10.1111/j.1751-9004.2009.00183.x>
- Pettigrew, T. F. (2021). *Contextual social psychology: Reanalyzing prejudice, voting, and intergroup contact*. American Psychological Association. <https://doi.org/10.1037/0000210-000>
- Phillips, J. E., & Olson, M. A. (2014). When implicitly and explicitly measured racial attitudes align: The roles of social desirability and thoughtful responding. *Basic and Applied Social Psychology*, 36(2), 125–132. <https://doi.org/10.1080/01973533.2014.881287>
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75(3), 811–832. <https://doi.org/10.1037/0022-3514.75.3.811>
- Scott, J. C. (2012). *Two cheers for anarchism*. Princeton University Press. <https://doi.org/10.1515/9781400844623>
- Shteynberg, G., Hirsh, J. B., Bentley, R. A., & Garthoff, J. (2020). Shared worlds and shared minds: A theory of collective learning and a psychology of common knowledge. *Psychological Review*, 127(5), 918–931. <https://doi.org/10.1037/rev0000200>
- Teachman, B. A., & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology*, 112(1), 100–109. <https://doi.org/10.1037/0021-843X.112.1.100>
- Tropp, L. R., & Barlow, F. K. (2018). Making advantaged racial groups care about inequality: Intergroup contact as a route to psychological investment. *Current Directions in Psychological Science*, 27(3), 194–199. <https://doi.org/10.1177/0963721417743282>
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107(1), 101–126. <https://doi.org/10.1037/0033-295X.107.1.101>