view Black people as violent, express greater racial resentment, and believe that anti-Black discrimination is an historical not contemporary problem (LeCount, 2017). Data from the real-world are thus congruent with those from the experimental paradigm that Cesario criticizes. As such, the bar for disqualifying experiments should be reasonably high, and calls to abolish this methodology should be greeted with healthy skepticism.

His case would also be more compelling if experimentalists failed to consider and contemplate boundary conditions, such as participant type (student vs. police), training/experience effects, cognitive load, and so on. Researchers not only study these nuances but also express clear caution and thoughtfulness. In their review, Payne and Correll (2020) conclude that "while an officer's performance on a laboratory task may provide valuable information, it cannot tell us whether race actually biases decisions about the use of force when police officers encounter suspects in the real world" (p. 36). Cesario's case that experiments create realities incongruent with the real world, and that central researchers extrapolate wildly from laboratory to the real world, are straw-man arguments. Similarly, his calls to consider the bigger context in police shootings would be compelling if he included the macro-level context, including its political and social structures, rather than his limited call to consider the specific micro-level situation (e.g., a specific shooter incident and its lead-up). He wants more information, but not too much.

Cesario's argument fits with a wider trend in academia to control the what-is-prejudice narrative and who gets to decide. As evidenced in the #BlackLivesMatter and #MeToo social movements, disadvantaged and marginalized groups are pleading for more voice at the table, not less. Psychologists express related concerns about the "extreme" and "overwhelming" Whiteness of psychology (see Dupree & Kraus, in press; Roberts, Bereket-Shavit, Dollins, Goldie, & Mortenson, 2020). In a culturally insensitive move, Cesario asks our discipline to direct more causal blame toward shooting victims and troubled children in classrooms, given their supposedly violent and undisciplined natures, for inviting their fates at the hands of the powerful.

As academics, we should be mindful that our ideas and work can be both used and misused. Defence attorneys for George Floyd's killing or the January 6th, 2021 Capitol Hill insurrection will appreciate the intellectual scaffolding these new academic trends offer to the Alt-Right, white supremacists, and those seeking to undo social change and justice. Our discipline lies at a critical crossroads; we can encourage epistemic inclusivity and incorporate more non-White voices, or we can become irrelevant (or detrimental) to the discipline of social studies.

Conflict of interest. No conflicts of interest.

References

Ba, B. A., Knox, D., Mummolo, J., & Rivera, R. (2021). The role of officer race and gender in police-civilian interactions in Chicago. *Science*, *371*(6530), 696–702. doi:10.1126/science.abd8694

Brandt, M. J., & Crawford, J. T. (2020). Worldview conflict and prejudice. *Advances in Experimental Social Psychology*, *61*, 1–66. https://doi.org/10.1016/bs.aesp.2019.09.002.

Dupree, C. H., & Kraus, M. W. (in press). Psychological science is not race neutral. *Perspectives in Psychological Science*. doi:10.1177/1745691620979820

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98–116.

Hangartner, D., Kopp, D., & Siegentaler, M. (2021). Monitoring hiring discrimination through online recruitment platforms. *Nature Human Behaviour*, *589*, 572–576. https://doi.org/10.1038/s41586-020-03136-0.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry*, *27*, 1–17. doi:10.1080/1047840X.2016.1082418

Hodson, G. (2021). Pushing back against the microaggression pushback in academic psychology: Reflections on a concept creep paradox. *Perspectives on Psychological Science*, *16*(5) 932–955. DOI: 10.1177/1745691621991863.

LeCount, R. J. (2017). More black than blue? Comparing the racial attitudes of police to citizens. *Sociological Forum*, *32* (S1), 1051–1072. doi:10.1111/socf.12367

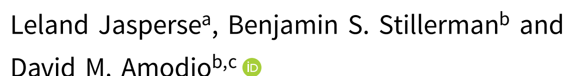
Payne, B. K., & Correll, J. (2020). Race, weapons, and the perception of threat. *Advances in Experimental Social Psychology*, *62*, 1–50. https://doi.org/10.1016/bs.aesp.2020.04.001.

Quillian, L., Lee, J. J., & Oliver, M. (2020). Evidence from field experiments in hiring shows substantial additional racial discrimination after the callback. *Social Forces*, *99*(2), 732–759. doi:10.1093/sf/soaa026

Quillian, L., Pager, D., Hexel, O., & Midtøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, *114*(41), 10870–10875. www.pnas.org/cgi/doi/10.1073/pnas.1706255114.

Roberts, S. O., Bereket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science*, *15*, 1295–1309. https://doi.org/10.1177/17456916209277.

Missing context from experimental studies amplifies, rather than negates, racial bias in the real world


Leland Jasperse[a], Benjamin S. Stillerman[b] and David M. Amodio[b,c]

[a]Department of English Language and Literature, University of Chicago, Chicago, IL 60637, USA; [b]Department of Psychology, New York University, New York, NY 10003, USA and [c]Department of Psychology, University of Amsterdam, 1001 NK Amsterdam, Netherlands.
ljasperse@uchicago.edu; ben.stillerman@gmail.com; david.amodio@nyu.edu, amodiolab.org

doi:10.1017/S0140525X21000868, e78

Abstract

We agree with Cesario's premise but reject his conclusion: Although experimental studies of racial stereotyping, weapons perception, and shoot decisions typically exclude real-world contextual factors and thus have limited relevance to race disparities (e.g., in policing), these excluded factors comprise systemic, institutional, and individual-level biases that are more likely to amplify racial disparities than negate them.

Cesario claims that experimental findings of racial bias are so disconnected from real-world situations that they "cannot and do not provide information about the nature of group disparities" (sect. 1, para. 2). Indeed, because such experiments are designed to isolate specific cognitive processes, they exclude myriad real-world factors that may otherwise influence intergroup behavior. However, we disagree with Cesario's conclusion that such factors overwhelm effects of social categories like race. In reality, the opposite is true: Real-world situations contain many layers of prejudice and discrimination, typically excluded from lab experiments, and these dramatically compound race effects.

Cesario argues that racial bias is only revealed in experiments when factors such as circumstantial information, group differences, and situational contingencies are omitted. Yet he all but ignores the many powerful layers of systemic, institutional, and

https://doi.org/10.1017/S0140525X21000017 Published online by Cambridge University Press

individual racism that pervade real-life interracial interactions. In fact, in U.S. policing, many of the situational factors omitted from lab studies are themselves shaped by race, such as racially motivated profiling and surveillance (Browne, 2015), stop-and-frisk policies (e.g., Cooper, 2018; Gelman, Fagan, & Kiss, 2007; Goel, Rao, & Shroff, 2016), and the use of discriminatory data-driven precision policing (Southerland, 2020). Although Cesario claims these real-world factors "overwhelm [the] strength of categorical bias" (Table 1 of target article), historical and sociological data suggest they actually *exacerbate* group disparities observed in experimental tasks.

To illustrate the supposedly race-neutralizing effect of real-world information, Cesario highlights a study by Correll, Wittenbrink, Park, Judd, and Goyle (2011) but misrepresents the finding. In this modified shooter task, targets are presented in either neutral or "dangerous, urban backgrounds" Cesario writes that the urban background – an instance of "missing information" reintroduced to a task – "completely eliminated racial bias in the decision to shoot" (sect. 4.1.1., para. 4). However, "dangerous, urban" settings are themselves racially coded from decades of segregationist housing policy, racist political rhetoric and media representations, and targeted over-policing (Gordon, 2020; Hurwitz & Peffley, 2005; Rhodes & Brown, 2019). Indeed, the data show that urban backgrounds actually increased the tendency to shoot White targets to the level of Black targets – an unsurprising effect given that these backgrounds themselves contain race-stereotypic cues.

As a real-world illustration, consider the NYPD's killing of Amadou Diallo, a case that galvanized research on implicit bias in shoot decisions: Four white NYPD officers patrolling the Bronx neighborhood of Soundview stopped Diallo, a young Black man "acting suspiciously" who allegedly matched the description of wanted criminal. When Diallo reached into his pocket for his wallet, the lead officer, per his testimony, misidentified it as a gun, triggering the group to shoot and kill Diallo. What other factors were at play that could have overwhelmed the subtle effect of automatic race associations? Notably, the officers were targeting a neighborhood that became majority-Black and over-policed following white flight, economic disinvestment, and redlining (Nonko, 2016; Stoudt, Fine, & Fox, 2011). Moreover, the officers were part of the infamous NYPD Street Crimes Unit, which expressly targeted *dangerous, urban* communities of color to turn up guns and drugs to meet quotas (Harring, 2000). Attributing Diallo's death to a quick decision made in ambiguous circumstances does leave out critical context from this scene, but this context amplifies disparities rather than ameliorates them (Amodio, 2015).

Although studies of implicit bias are often inspired by real-world incidents, they are rarely (if ever) designed to explain them. Instead, they aim to isolate and illuminate basic mechanisms of race processing in the mind; asking, for example, *Can race influence automatic thought and quick decisions?* Such experiments are rarely presented as complete accounts of real-world disparities and expressions of prejudice. Curiously, the article Cesario singles out as "a prototypical example" (sect. 2, para. 2) of this practice, by Moss-Racusin, Dovidio, Brescoll, Graham, and Handelsman (2012), is a field study on gender bias in job applicant evaluations that uses none of the methods he critiques. Moreover, social psychologists have long considered the roles of additional information, forces, and contingencies as moderators of category-based stereotyping (e.g., Amodio & Swencionis, 2018; Darley & Gross, 1983; Dovidio & Gaertner, 2000; Fiske & Neuberg, 1990). The deficiencies Cesario attributes to social psychology appear to concern its translation more than the science itself.

We see a different concern with reductionist experimental studies which, we believe, is much more pressing (Jasperse & Stillerman, 2021): By presenting racial bias as a subtle, unintentional spandrel of the mind, these studies problematically reduce the broad, structural nature of racism to a transient impulse. Consequently, they misdirect efforts toward ineffective training programs (Worden et al., 2020) and give cover to the more pernicious effects of systemic, institutional, and blatant racism. Hence, in addition to underestimating the magnitude of bias, such studies draw attention away from its deeper causes.

Finally, we feel compelled to comment on the selective scholarship and rhetoric in this target article. Cesario elides evidence that racial bias is a pervasive dimension of policing and criminal justice – one that inflects (and exceeds) moment-to-moment individual cognition. He then suggests that observed real-world disparities are due mainly to behavioral differences between groups. For example, he argues that racial disparities in policing may be more a product of different racial groups' criminal tendencies than bias on the part of police officers. Although he hastens to "make no claims about the origin of these group differences" (sect. 1, para. 6), a casual reader could be forgiven for thinking that Cesario believes elevated criminality "might very well be" (sect. 1, para. 6) a trait feature of racial minorities. This rhetorical pattern – to deny the severity of racial bias and then suggestively attribute disparities to individual merits of group members – follows a familiar refrain known to social psychologists as *modern racism*. Regardless of the authors views and intentions, it is concerning to see this device in mainstream scientific discourse.

In summary, we accept Cesario's premise but reject his conclusion; the many real-world factors often missing from sociocognitive experiments of racial bias are themselves the product of systemic, institutional, and individual racism. To the extent real-world factors overwhelm experimentally observed patterns of bias, the effect of racism is likely much stronger.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

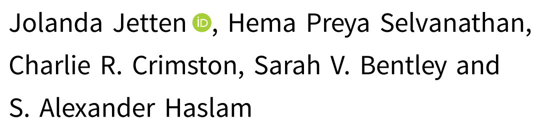
Conflict of interest. None.

References

Amodio, D. M. (2015). Prejudiced? Me? *New Scientist 227*:26–27.
Amodio, D. M., & Swencionis, J. K. (2018). Proactive control of implicit bias: A theoretical model and implications for behavior change. *Journal of Personality and Social Psychology 115*(2):255–275.
Browne, S. (2015). *Dark matters*. Duke University Press.
Cooper, F. R. (2018). A genealogy of programmatic stop and frisk: The discourse-to-practice circuit. *University of Miami Law Review 73*(1), 1–78.
Correll, J., Wittenbrink, B., Park, B., Judd, C. M., & Goyle, A. (2011). Dangerous enough: Moderating racial bias with contextual threat cues. *Journal of Experimental Social Psychology 47*:184–189.
Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology 44*(1):20–33.
Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science 11*(4):315–319.
Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology 23*:1–74.
Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York city police department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association 102*(479), 813–823.
Goel, S., Rao, J. M., & Shroff, R. (2016). Precinct or prejudice? Understanding racial disparities in New York city's stop-and-frisk policy. *The Annals of Applied Statistics 10* (1):365–394.

Gordon, D. (2020). The police as place-consolidators: The organizational amplification of urban inequality. *Law & Social Inquiry 45*(1):1–27.
Harring, S. L. (2000). The Diallo verdict: Another "tragic accident" in New York's war on street crime? *Social Justice 27*(1), 9–18.
Hurwitz, J., & Peffley, M. (2005). Playing the race card in the post-Willie Horton era: The impact of racialized code words on support for punitive crime policy. *Public Opinion Quarterly 69*(1):99–112.
Jasperse, L., & Stillerman, B. (2021). Beyond bias: The case for an abolitionist psychology. Los Angeles Review of Books.
Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences 109*(41):16474–16479.
Nonko, E. (2016). Redlining: How one racist, depression-era policy still shapes New York real estate. Brick Underground.
Rhodes, J., & Brown, L. (2019). The rise and fall of the "inner city": Race, space and urban policy in postwar England. *Journal of Ethnic & Migration Studies 45*(17):3243–3259.
Southerland, V. (2020). The intersection of race and algorithmic tools in the criminal legal system. *Maryland Law Review 80*(3), 487–566.
Stoudt, B. G., Fine, M., & Fox, M. (2011). Growing up policed in the age of aggressive policing policies. *New York Law School Law Review 56*:1331–1370.
Worden, R. E., McLean, S. J., Engel, R. S., Cochran, H., Corsaro, N., Reynolds, D., … Isaza, G. T. (2020). The impacts of implicit bias awareness training in the NYPD.

Experiments make a good breakfast, but a poor supper



Jolanda Jetten, Hema Preya Selvanathan, Charlie R. Crimston, Sarah V. Bentley and S. Alexander Haslam

School of Psychology, The University of Queensland, St. Lucia, 4072 QLD, Australia.
j.jetten@psy.uq.edu.au; https://psychology.uq.edu.au/profile/2317/jolanda-jetten; h.selvanathan@uq.edu.au; https://psychology.uq.edu.au/profile/7410/hema-preya-selvanathan; c.crimston@uq.edu.au; https://psychology.uq.edu.au/profile/2698/charlie-crimston; s.bentley@uq.edu.au; https://psychology.uq.edu.au/profile/2536/sarah-bentley; a.haslam@uq.edu.au; https://psychology.uq.edu.au/profile/3181/alex-haslam

doi:10.1017/S0140525X21000662, e79

Abstract

Cesario's analysis has three key flaws. First, the focus on whether an effect is "real" (an "effects flaw") overlooks the importance of theory testing. Second, obsession with effects (a "fetishization flaw") sidelines theoretically informed questions about when and why an effect may arise. Third, failure to take stock of cultural and historical context (a "decontextualization flaw") strips findings of meaning.

Cesario provides a number of good reasons why we should be cautious about relying solely on experimental findings to understand the social world around us. While we welcome the focus on experimental validity (after years of focusing more or less exclusively on problems associated with replication and reliability), unfortunately, his own analysis falls foul of some of the problems that it seeks to rectify. There are three specific flaws in his reasoning, and all three are commonly observed in researchers' understanding of what experiments are meant to do and how they should be used.

First, Cesario's analysis misunderstands the purpose of experiments. Their function is not to try as hard as possible to mimic aspects of the world outside the laboratory so that researchers can establish whether a given effect is observable in the world and hence "real" (e.g., whether or not police officers are racially biased). To imagine that they are is to fall prey to an "effects flaw" in which experimental outcomes are privileged over the processes that produce them.

Instead, then, experiments and the evidence they produce are better suited to the task of testing *theories* of human psychology and behaviour. They do this principally by helping us to understand under *what conditions* a given effect is observed, and *what mechanisms* underlie that effect. Indeed, by focusing on effects rather than processes, Cesario's analysis fails to capitalise on the key value of experiments – namely their capacity to support theory development (Haslam & McGarty, 2001; Swann & Jetten, 2017).

This "effects flaw" is not just present in Cesario's analysis, but is a pervasive problem in the social psychological literature. It is perhaps most apparent in reports of the classic studies in social psychology (e.g., Milgram's obedience studies and Zimbardo's Stanford Prison Experiment; see Smith & Haslam, 2017). For instance, because of the "effect flaw" the contribution of Milgram's obedience studies is routinely misunderstood. For the real theoretical value of the work can be seen to lie less in the 65% obedience rate that was observed in the so-called "baseline condition" (the classic effect reported in most textbooks) than in the many variants that Milgram conducted to explore the conditions under which obedience is either far greater or far weaker (see Jetten & Mols, 2014; Reicher, Haslam, & Smith, 2012). To be sure, experimental effects can capture our attention and make the case for much-needed theory development, but without a theoretical focus and grounding, their contribution is unproductively circumscribed.

Second, while we agree that, on its own, experimental evidence is of limited use, we argue that what is needed is a proper analysis of how experimental evidence should be complemented with other forms of evidence. Here, we would argue that experimental evidence should never be considered in isolation, but always in conjunction with data sourced using complementary methods (e.g., field surveys, longitudinal research, and qualitative work). What is more, theory-derived hypotheses need to be examined in a range of different contexts. Unfortunately, although, experimental evidence is too often seen as the "gold (and only) standard" for our field, with evidence gleaned via other means relegated to the margins.

This prioritization of experimental effects contributes to a "fetishization flaw" associated with what Reicher (2000) refers to as methodolatry. As a result of this there is little incentive for researchers to move out of the lab, and once an "effect" is established within a controlled laboratory setting, it hardly ever comes out of it. The experimental paradigm, therefore, becomes equated with the phenomena itself. This exacerbates the consequences of the first flaw by cultivating an obsession with (the replication of) experimental effects and attendant neglect of broader questions of process. In short, questions of "when" and "why" are crowded out by questions of "whether" and "how much" in ways that stymie and suppress theory development and the deep understanding that accompanies it. As the replication crisis of recent years attests, this narrowing of the field has not served social psychology well.

Third, alongside these issues, a "decontextualization flaw" means that researchers typically use experiments for hypothetico-deductive purposes in a quest to discover "objective truth." This epistemology generally assumes value neutrality and context independence and tends to catalogue psychological effects with scant regards to the broader historical and societal contexts in which they arise (Adams, Estrada-Villalta, Sullivan, & Markus, 2019).

In crucial ways, this has led to the disappearance of the "social" in social psychology (see Greenwood, 2003). For it is important to