Research Article

Psychological Science
2023, Vol. 34(4) 424–434

DOI: 10.1177/09567976221150617
www.psychologicalscience.org/PS

The Impact of Implicit-Bias-Oriented Diversity Training on Police Officers' Beliefs, Motivations, and Actions

Calvin K. Lai[1] and Jaclyn A. Lisnek[2]
[1]Department of Psychological & Brain Sciences, Washington University in St. Louis, and [2]Department of Psychology, University of Virginia

Abstract

U.S. police departments have attempted to address racial inequities in policing with diversity training. However, little research has evaluated whether these trainings are effective at changing officers' beliefs, motivations, and actions. To examine their efficacy, we tested a day-long implicit-bias-oriented diversity training designed to increase U.S. police officers' knowledge of biases, concerns about bias, and use of evidence-based strategies to mitigate bias (total N = 3,764). The training was immediately effective at increasing knowledge about bias, concerns about bias, and intentions to address bias, relative to baseline. However, the effects were fleeting. Although the training was linked to higher knowledge for at least 1 month, it was ineffective at durably increasing concerns or strategy use. These findings suggest that diversity trainings as they are currently practiced are unlikely to change police behavior. We conclude with theorizing about what organizations and training programs could do for greater impact.

Keywords

diversity training, intervention, policing, bias, race, prejudice, open data, open materials, preregistered

Received 1/20/22; Revision accepted 12/20/22

At almost any point of contact between police officers and Black individuals in the United States, there is evidence of unequal treatment. U.S. police officers are more likely to stop, question, arrest, injure, or kill Black people than White people (Glaser, 2015). Subtle or implicit biases have been identified as a psychological mechanism underlying these inequities in police behavior (Spencer et al., 2016). These implicit biases may cause police officers to perceive Black people as more dangerous than is warranted. For instance, a laboratory experiment found that police recruits are more likely to shoot unarmed Black men than unarmed White men in simulated ambiguous scenarios (Ma et al., 2013). Police officers are armed by the state to ensure public safety, and yet their actions often fall short of doing so equitably.

Recognizing these disparities, police departments across the United States have stepped up their efforts to reduce racial inequities with bias-training programs (President's Task Force on 21st Century Policing, 2015). Of particular interest is the rise of implicit-bias-oriented diversity training programs that educate police about implicit bias and provide strategies to mitigate the impact of that bias on behavior. A survey of 109 U.S. police departments in large metropolitan areas indicated that 96% have some form of implicit-bias-oriented diversity training program (CBS News, 2019). Despite these widespread efforts, most bias-training programs have not been evaluated empirically and are not informed by psychological research on bias reduction or lasting behavior change (James, 2017; Paluck et al., 2021).

The little research on the efficacy of modern diversity training programs has been almost exclusively conducted in non-law-enforcement settings and may not generalize to policing (Devine & Ash, 2022; Dobbin et al., 2015; Paluck et al., 2021). As the face of governmental social

Corresponding Author:
Calvin K. Lai, Department of Psychological & Brain Sciences, Washington University in St. Louis
Email: calvinlai@wustl.edu

control (Soss & Weaver, 2017), police officers hold different ideological commitments than the general population, are beholden to distinct hierarchical institutional structures, and must meet job demands that are unlike those of almost any other profession. For instance, U.S. police officers are socialized into a "warrior" worldview that is preoccupied with dominance, violence, and maintaining officer safety to contend with the potential danger that comes with the job (Sierra-Arévalo, 2021; Stoughton, 2014). This hypervigilance takes precedence over cultivating community relationships and may stymie efforts to promote fair and equitable treatment on the job. Understanding the efficacy of diversity training on police officers, then, requires direct assessment.

We assessed the Managing Bias program, which was developed by the Anti-Defamation League to reduce the influence of bias in interactions and decision-making by law enforcement, contribute to improved police–community relations, and increase officer safety. In the Managing Bias program, pairs of educators jointly led an interactive day-long workshop that emphasized discussion and active learning over lecturing. Throughout the workshop, officers participated in activities that raised their awareness of bias. These activities educated officers about the definition and origins of bias, the influence of culture and identity on social experience, and gaps in understanding between police officers and community members. By the end of this section, officers were expected to understand the differences between implicit and explicit bias, identify how their worldview has been shaped by their culture and identity, and appreciate how biases may affect behavior.

The final section of the training program focused on building skills to manage bias in policing. Officers were taught five practices that had been shown to reduce biased behavior. The first strategy was mindfulness: intentionally bringing awareness to the present moment (Lueke & Gibson, 2015). The other four strategies were adapted from Devine and colleagues' (2012) prejudice-habit-breaking intervention, which has been linked to enduring increases in concern about discrimination and improvements in workplace culture in university faculty and student samples (Carnes et al., 2015; Forscher et al., 2017). The four strategies were stereotype substitution, perspective-taking, individuation, and diversity exposure. Stereotype substitution involved replacing thoughts of negative stereotypes with positive mental images (Monteith, 1993), perspective-taking involved taking the perspective of another person to understand their point of view (Galinsky & Moskowitz, 2000), and individuation involved getting to know people as unique individuals (Rubinstein et al., 2018). Finally, officers were instructed to actively seek opportunities to learn about people from different backgrounds through direct and indirect intergroup contact (Pettigrew & Tropp, 2006). By the end of this section, officers were expected to be able to identify thoughts and actions linked to implicit bias and to manage their impact on behavior.[1]

Statement of Relevance

Enduring racial disparities in policing and a decline in trust of police officers has spurred U.S. law-enforcement agencies to invest in diversity training. However, little is known about whether diversity trainings change what officers believe or do. We tested a day-long implicit-bias-oriented diversity training on 3,764 U.S. police officers that was developed by a premier nongovernmental organization. Officers who took the training were more knowledgeable about bias and more motivated to address bias at work, relative to baseline. However, those effects did not last. These findings suggest that diversity trainings as they are currently practiced are unlikely to reduce racial inequity in policing. In future training efforts, agencies may consider strong integration with organizational initiatives, longer periods of training, or different approaches.

Open Practices Statement

All data, analysis scripts, and survey materials have been made publicly available at OSF and can be accessed at https://osf.io/vfdrt/. Analyses for both cohorts were preregistered at https://osf.io/d29ps and https://osf.io/6k3fg.

Method

To assess the training, we conducted a study of several thousand sworn officers of police departments within the United States who took part in training sessions that used the Managing Bias curriculum. To preserve the confidentiality of the police department partners that we worked with for this study, we do not name the department or departments that we worked with. The police departments had a history of Black–White racial disparities in police-initiated stops and use of force. In total, we evaluated 251 training sessions that were conducted by 24 different educators. The size and diversity of our sample allowed us to assess the general efficacy of the Managing Bias training and also afforded opportunities to examine the characteristics of police officers and educators that were linked to changes in beliefs, motivations, and actions.

Participants

Participants were sworn officers of a police department or departments within the United States who took part in one of 251 training sessions during either one of two data-collection waves (July 2019 to March 2020; September 2020 to January 2021). We collected data from two cohorts because of the COVID-19 pandemic. Data for the first cohort were collected before the COVID-19 pandemic, and this training was conducted completely in person. Surveys were administered immediately before training and immediately after training. Data for the second cohort were collected during the COVID-19 pandemic. We adapted the training for remote learning, with about half of the training taking place using self-paced modules and half of the training taking place synchronously with two educators over a video conferencing platform. Surveys were administered immediately before the first training session, immediately after the second training session, and 1 month after the second training session. Officers were mandated to take the training as part of their work activities but voluntarily participated in the study. The study was approved by an institutional review board at Washington University in St. Louis. We planned to collect data from as many officers as possible during our data-collection windows.

In both cohorts, data were excluded if the officer had completed only a single survey in the multisurvey design or if we failed to match surveys between time points. Because the second cohort's data were collected online instead of via pen and paper, we also preregistered exclusions for providing duplicate surveys, completing the survey too quickly, or accessing the study on a date that was hypothetically impossible (e.g., a baseline survey completed after training was completed). See the Supplemental Material available online for a complete description of exclusions.

Combined, we obtained usable data from 3,764 officers (after exclusions) who participated in the first two surveys. Because the third survey that was delivered 1 month after was administered only to the second cohort, we ended up with smaller samples for those analyses. We obtained usable data from 173 participants who completed the first and third surveys, 73 officers who completed the second and third surveys, and 53 participants who completed all surveys.

We encountered different challenges with participant retention in each cohort. In the first cohort, 4,659 officers were trained in person. We obtained very high rates of pen-and-paper survey participation immediately before training (4,177; 91%) and immediately after training (4,426; 97%). After exclusions, we obtained usable data from 3,417 (73%) officers trained in the first cohort who completed both surveys. In the second cohort, 3,103 officers were trained online. As the survey could be accessed multiple times and there were technical issues that led to duplicate surveys and evidence of survey misbehavior (see the Supplemental Material), we obtained 4,125 completed surveys immediately before training. After exclusions, we obtained 2,883 usable surveys from immediately before the first training session. The other two survey administrations (immediately after and 1 month after training) had minimal evidence of technical issues or survey misbehavior but lower completion rates. We collected 647 completed surveys immediately after training, of which 620 (96%) were retained after exclusions. We also collected 318 completed surveys 1 month after training, of which 300 (94%) were retained after exclusions. Because of low retention and the use of a method to create identifier codes without collecting directly identifying data that leads to low matching rates (Yurek et al., 2008), we failed to match many of the surveys in the second cohort. In exploratory analyses of both cohorts, we also found significant evidence of differential attrition across 11 of 68 analyses, which we describe more in the Supplemental Material.

Participants had worked an average of 15 years in their departments, with 77% having worked in their department for 5 years or more. Eighty-six percent of participants were below sergeant rank, 77% were male, and 64% had a bachelor's degree or higher. Out of 3,104 participants who reported their race, 47% were White, 20% were Black, 27% were Hispanic/Latino, 2% were Asian, and 3% were multiracial. In their daily work, officers reported frequently interacting with White people 62% of the time; Black people 90% of the time; Hispanic or Latino people 70% of the time; Asian people 24% of the time; transgender people 13% of the time; and gay, lesbian, or bisexual people 34% of the time.

Procedure

Officers completed up to three surveys that assessed self-reported beliefs, attitudes, motivations, and behavior. The first survey was a baseline survey administered to officers before the training began. This survey assessed general understanding and concern about bias, baseline usage of strategies to manage bias, and several baseline characteristics that are relevant to police training (i.e., police identity centrality, expectations of respect from community members, and demographic information). The second survey assessed immediate effectiveness and was administered immediately after the end of training. This survey assessed officers' general understanding and concern about bias, intentions and motivations to use the strategies they

learned to manage bias, and perceived understanding and perceived efficacy of those strategies. This survey was also supplemented by open-ended written questions asking officers to describe their reactions to the training (see the Supplemental Material). Finally, the third survey was administered 1 month after training. It was almost identical to the second survey, except that it measured reported usage of strategies instead of intentions to use strategies and removed a filler question. The third survey was administered only during the data collection for the second cohort, which meant that much less data were available for this survey than for the first two surveys.

Measures

Knowledge of bias (all surveys). Assessment of understanding of key concepts was tailored to the concepts covered in the Managing Bias training program. Participants were asked to report their agreement with three statements: "Subtle or implicit biases influence my decision making about other people"; "Whether I am aware of it or not, I use a person's race or ethnicity to form an impression of the kind of person they are"; and "My cultural background influences how I perceive other people." Agreement was measured on a 7-point scale ranging from 1 (*Strongly disagree*) to 7 (*Strongly agree*), with higher scores indicating a greater understanding of bias (αs = .81, .86, and .90 at the first, second, and third surveys, respectively).

Concern about bias (all surveys). We adapted the Bias Awareness Scale (Perry et al., 2015) to assess concern about unintentional bias toward five minority groups. Participants reported their agreement with the following statement: "I worry that I act in an unintentionally biased way toward [group]." The six groups assessed were Black people; Asian people; Hispanic or Latino people; transgender people; and gay, lesbian, or bisexual people. Agreement was measured on a 7-point scale ranging from 1 (*Strongly disagree*) to 7 (*Strongly agree*), with higher scores indicating greater worry about bias (αs = .98, .98, and .90 at the first, second, and third surveys, respectively). We also included a filler question about White people that was not used in data analysis.

Beliefs about the malleability of bias (all surveys). Prior research suggests that people who believe bias to be more malleable rather than fixed are more likely to act positively in interracial contexts (Rattan & Georgeac, 2017). If learning about implicit biases degrades a sense of control over biases, then implicit bias education could backfire. To examine this possibility, we measured lay beliefs about the malleability of bias using two questions adapted from the Theories of Prejudice Scale that assessed agreement with the following statements: "People have a certain amount of bias toward other people and they can't really change that" (reverse scored) and "People's level of bias toward other people is something very basic about them that they can't change very much" (reverse scored; Carr et al., 2012). Agreement was measured on a 7-point scale ranging from 1 (*Strongly disagree*) to 7 (*Strongly agree*), with higher scores indicating greater beliefs that bias was malleable rather than fixed (αs = .83, .82, and .66 at the first, second, and third surveys, respectively).

Use of strategies to manage bias (first and third surveys). We gave officers brief descriptions of each of the five key strategies to manage bias. For each of the strategies, we then asked them, "Over the past seven days, how many times have you used this strategy?" Responses were made on a 7-point scale from 0 (*0 times*) to 6 (*6 or more times*; αs = .83 and .90 at the first and third surveys, respectively). These questions were adapted from Devine and colleagues' (2012) study.

Intention to use strategies (second survey). We gave officers brief descriptions of each of the five key strategies to manage bias. For each of the strategies, we then asked them, "Over the next seven days, how many times do you plan to use this strategy?" Responses were made on a 7-point scale from 0 (*0 times*) to 6 (*6 or more times*; α = .96). These questions were adapted from Devine and colleagues' (2012) study.

Other survey measures. We also collect additional measures of perceptions of the strategies to manage bias and baseline characteristics of police officers, such as police identity centrality, interracial contact, and expectations of respect from community members. See the Supplemental Material for a detailed description of these measures.

Results

We made several deviations from the preregistration for reporting purposes. First, because results were consistent across both cohorts (with several minor exceptions), we used a data set combining both cohorts for our reported analyses that compare the first and second surveys. Second, our preregistered analyses of the second cohort were done using several different approaches to assess robustness to missing data. The results were generally consistent across approaches, so we report the analyses with the largest effective samples (n = 173 for analyses comparing the first and third surveys; n = 74 for analyses comparing the second and third surveys). Third, there were several analyses that were preregistered for the second cohort but not the first:

analyses involving characteristics of educators and analyses using mixed models instead of repeated measures analyses of variance (ANOVAs). Analyses involving mixed models were consistent with repeated measures ANOVAs. The Supplemental Material describes all preregistered analyses that are not reported in the main text. We report Cohen's *d*s as standardized effect sizes for mixed models. This statistic reflects the mean difference between time points divided by the standard deviation of the difference between time points.

Before the training program, officers expressed low understanding and concern about bias. On average, officers did not believe in the existence of subtle or hidden biases that influence their behavior, $M = 3.09$ out of 7, $SD = 1.53$, and were generally unconcerned about it, $M = 2.24$ out of 7, $SD = 1.46$. There was considerable variation in how often officers used the strategies taught in the Managing Bias training program at baseline. Officers were most likely to engage in perspective-taking and mindfulness exercises; more than 70% of officers reported using those strategies at least once in the past week. In contrast, officers were less likely to report using stereotype substitution (41%), diversity exposure (52%), or individuation (37%) strategies over the same time period.

We found that the training was immediately effective at increasing knowledge about bias relative to baseline. As can be seen in Figure 1, officers were more likely to report believing in the existence of subtle or hidden biases after taking the training, $b = 0.69$, $SE = 0.02$, $t(3673.21) = 28.51$, $p < .001$, $d = 0.48$. The immediate increase in knowledge of bias was durable and persisted several weeks after the training. One month after training, officers continued to be more likely to report believing in the existence of subtle or hidden biases relative to baseline, $b = 0.46$, $SE = 0.12$, $t(172) = 3.89$, $p < .001$, $d = 0.30$, and did not significantly differ from their knowledge immediately after training, $b = 0.06$, $SE = 0.20$, $t(73.47) = 0.29$, $p = .77$, $d = 0.03$. When officers were asked to describe their thoughts about the training, many reported that it was surprising and insightful. For instance, one officer wrote, "It has opened my eyes to the bias we all have as human beings" and another said, "I really liked the course because it opened my eyes to implicit biases I never knew I had."

The training temporarily increased concern about bias, which may be an important prerequisite to egalitarian action (Perry et al., 2015). After the training, officers expressed more concern about unintended biases relative to baseline, $b = 0.39$, $SE = 0.02$, $t(3715.43) = 17.04$, $p < .001$, $d = 0.28$. However, after 1 month, officers' concerns about bias returned to baseline levels, $b = -0.06$, $SE = 0.11$, $t(172) = -0.61$, $p = .55$, $d = -0.05$.

We also examined the possibility that training may have reduced beliefs in the malleability of bias, which may disempower trainees from pursuing action (Carr et al., 2012; Rattan & Georgeac, 2017). We found that the training temporarily reduced beliefs in the malleability of bias, $b = -0.12$, $SE = 0.03$, $t(3714.66) = -4.45$, $p < .001$, $d = -0.07$, indicating that immediate increases in knowledge and concern about bias were slightly offset by perceptions that bias was less malleable. However, these effects were short-lived. After 1 month, officers' perceptions of the malleability of bias did not significantly differ from their perceptions at baseline, $b = -0.18$, $SE = 0.13$, $t(172) = -1.45$, $p = .15$, $d = -0.11$, and were significantly higher compared with immediately after training, $b = 0.53$, $SE = 0.18$, $t(73.37) = 2.97$, $p = .004$, $d = 0.35$.

Officers were empowered and motivated to use the five strategies to manage bias that were taught in the training. Officers expressed confidence in their understanding of the strategies, $M = 4.19$, $SD = 0.89$; believed the strategies to be effective for addressing bias, $M = 3.83$, $SD = 1.00$; and were motivated to use the strategies, $M = 3.68$, $SD = 1.07$. Officers also reported much higher intentions to use the five strategies over the next week relative to their self-reported baseline usage, $b = 1.37$, $SE = 0.03$, $t(3571.42) = 45.23$, $p < .001$, $d = 0.78$. Officers' qualitative feedback on the training supported these findings. One officer said that they "will definitely use [these tools] in the future," and another said it was "great information that I will utilize for the rest of my career." Officers intended to use each of the five strategies an average of 3.25 times each in the next week. Variations in the intended use of each of the five strategies were much smaller than variations in the reported baseline usage of the five strategies. The largest gains were in stereotype substitution, the strategy that was used least at baseline, $b = 1.91$, $SE = 0.04$, $t(3629.44) = 50.66$, $p < .001$, $d = 0.85$. The smallest gains were in the strategies that were used most at baseline, perspective-taking, $b = 0.84$, $SE = 0.04$, $t(3639.99) = 23.50$, $p < .001$, $d = 0.40$, and mindfulness, $b = 0.85$, $SE = 0.04$, $t(3638.57) = 21.30$, $p < .001$, $d = 0.36$.

However, officers did not follow through on their intentions to use the strategies 1 month later. As can be seen in Figure 2, officers reported lower usage of strategies compared with the intentions they reported immediately after training, $b = -1.46$, $SE = 0.20$, $t(73.67) = -7.19$, $p < .001$, $d = -0.90$. Unexpectedly, officers also reported less use of strategies 1 month later compared with baseline, $b = -0.23$, $SE = 0.06$, $t(172) = -3.60$, $p < .001$, $d = -0.27$. They continued to think the strategies were feasible, $M = 3.92$, $SD = 1.06$, $t(172) = 36.26$, $p < .001$, $d = 3.70$, and did not perceive them as significantly less feasible compared with immediately after training,

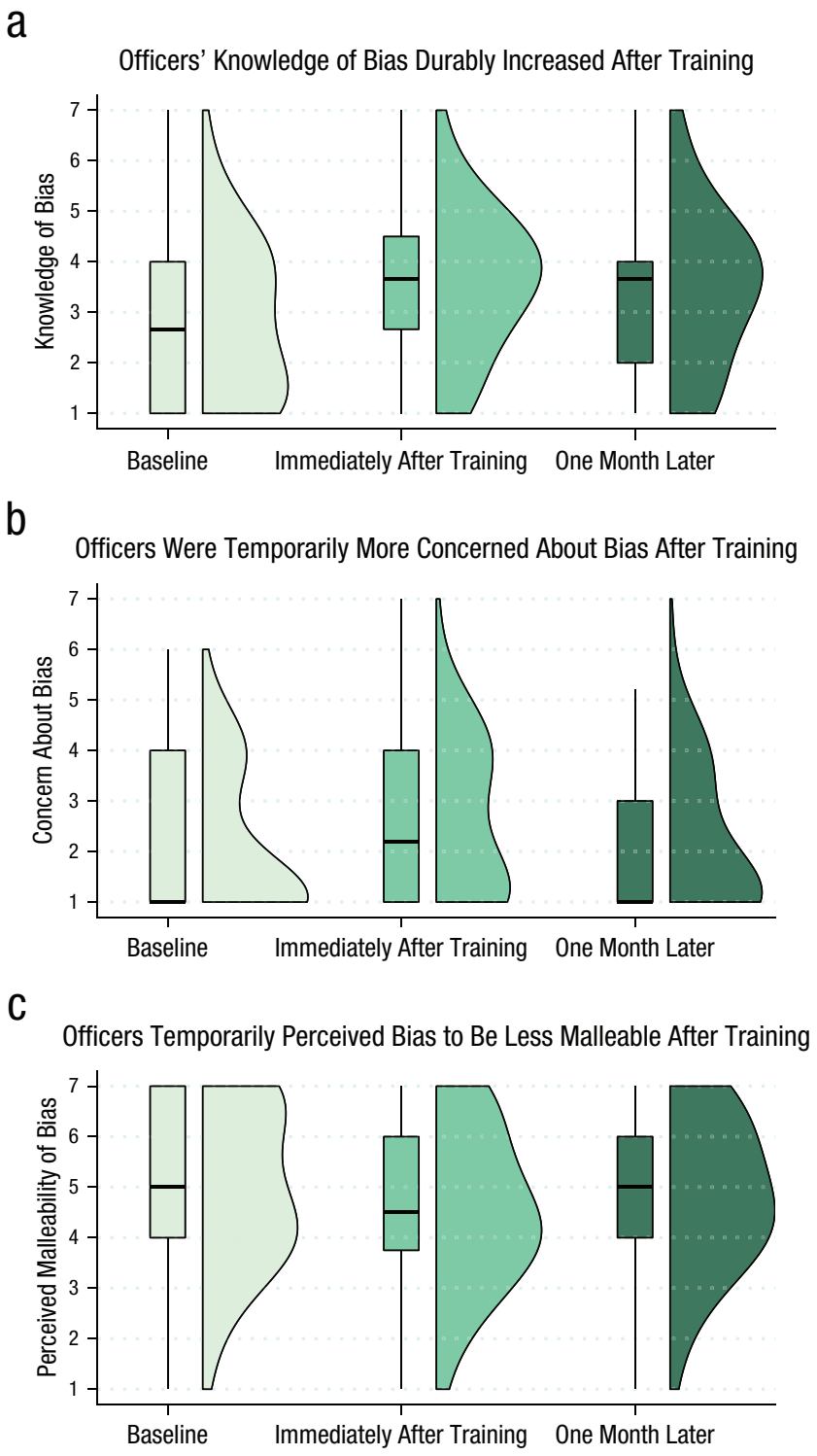

Fig. 1. Officers' knowledge about bias (a), concerns about bias (b), and perceptions of the malleability of bias (c) before, immediately after, and 1 month after training. Data are from officers in the second cohort who took the baseline survey. The line inside each box represents the median, the box boundaries represent the lower 25th percentile and upper 75th percentile, and the error bars indicate 1.5 times the interquartile range. The half-violin plots show the probability density of the data.

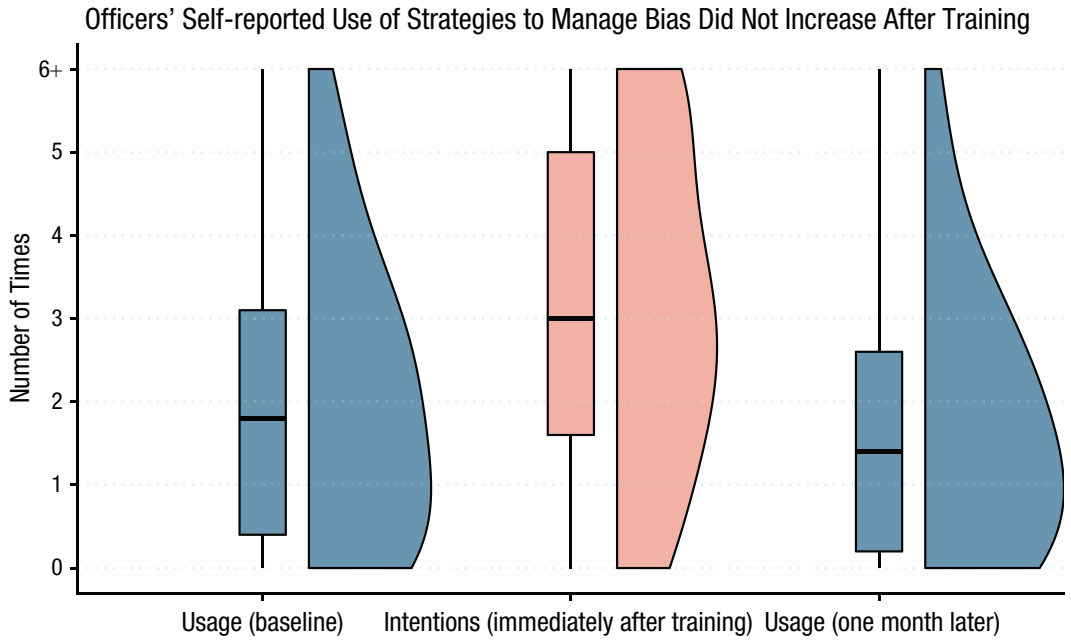

Fig. 2. Officers' self-reported use of strategies (baseline, 1 month later) and intentions to use strategies (immediately after training) to manage bias. Data are from officers in the second cohort who took the baseline survey. The line inside each box represents the median, the box boundaries represent the lower 25th percentile and upper 75th percentile, and the error bars indicate 1.5 times the interquartile range. The half-violin plots show the probability density of the data.

$b = -0.09$, $SE = 0.10$, $t(73.95) = -0.95$, $p = .35$, $d = -0.12$. However, they were less motivated to use the strategies, $b = -0.28$, $SE = 0.12$, $t(73.70) = -2.35$, $p = .021$, $d = -0.30$, and thought they would be less effective for mitigating bias, $b = -0.32$, $SE = 0.11$, $t(73.64) = -2.82$, $p = .0061$, $d = -0.35$. The most straightforward interpretation is that officers were less likely to engage in the strategies. Perhaps officers used (or considered using) the strategies and did not find them helpful for their daily work. Officers may have also shifted to a more conservative interpretation of what counts as using the strategies effectively after learning about effective strategy use in the training. Engaging in effective perspective-taking, for example, might mean something more intensive than what they had originally believed at baseline. Officers may have also reported lower strategy use to express recognition of a failure to meet their intentions. On average, officers reported intentions to use strategies that were almost double what they reported doing a month later.

We examined whether immediate changes in our key outcomes were related to characteristics of police officers. We examined three demographic factors (i.e., race/ethnicity, gender, level of education), five social-psychological individual differences (i.e., police identity centrality, expectations of respect from community members, baseline knowledge, baseline concern about bias, baseline beliefs in the malleability of bias), and three factors about their work experience (i.e., contact with minority groups, rank, years in department). Many of these characteristics were related to baseline differences in our key variables. However, characteristics of the officers generally did not predict the amount of learning, with several exceptions that were not robust to changes in model specification. See the Supplemental Material for a summary of the exceptions.

We also examined whether immediate changes in our key outcomes were related to characteristics of educators. The 24 educators in our studies were diverse in race (54% Black, 38% White, 4% Asian, 4% multiracial), gender (63% female, 37% male), and age (range: 30s to 60s). One Black and one non-Black educator were assigned to each training session when possible (83% Black–White pairs), and educator pairs tended to be mixed gender (61% female–male pairs). Educators also varied in life experience and profession, including lawyers, mediators, teachers, and diversity consultants. One of the educators was an ex-law-enforcement professional, and many educators had a close relationship with a law-enforcement professional (e.g., parent, spouse; 38% of sessions).

We found that some characteristics of the educators were linked to outcomes when the training was in person but not when the training was online. This was the only set of reliable differences we found between in-person and online trainings. When the training was in

person, sessions with male–male pairs of educators were linked to lower gains in knowledge about bias than sessions with mixed-gender pairs, $b = 0.18$, $SE = 0.07$, $t(3290.17) = 2.42$, $p = .016$, or female–female pairs, $b = 0.18$, $SE = 0.08$, $t(3294.48) = 2.18$, $p = .030$. Officers in sessions with male–male pairs also showed lower immediate increases in concern than officers in sessions with mixed-gender pairs, $b = 0.16$, $SE = 0.07$, $t(3342.44) = 2.29$, $p = .022$. Sessions with the one ex-law-enforcement educator showed larger gains in knowledge, $b = 0.39$, $SE = 0.08$, $t(3314.74) = 4.84$, $p < .001$; concern, $b = 0.20$, $SE = 0.08$, $t(3364.12) = 2.61$, $p = .009$; and intended strategy use relative to baseline usage, $b = 0.38$, $SE = 0.10$, $t(3199.88) = 3.80$, $p < .001$, than sessions without that educator during in-person training. However, this result should be interpreted with caution because sessions with that one educator may have differed from other sessions in many other ways. Educator race and relationships with law-enforcement professionals were not linked to learning outcomes.

Discussion

We investigated the impact of an implicit-bias-oriented diversity training program on U.S. police officers. The training was effective at durably increasing its most proximal outcome: knowledge of bias. In contrast, the training only temporarily increased concerns about bias and motivation to use strategies. A temporary increase in concern and motivation may have been unproblematic if the increase spurred the development of behavioral habits that persisted long after training (Frey & Rogers, 2014). However, self-reported strategy use did not increase 1 month afterward. Educating officers about implicit bias was effective for durably raising their awareness about the existence of subtle or implicit biases but little else.

The day-long Managing Bias training represented a best-case scenario for diversity trainings in society today. It was developed by the Anti-Defamation League, a premier nongovernmental organization for combating discrimination, and delivered by a diverse range of 24 educators. The day-long training was also more intensive than other diversity trainings, which are often only 1 to 3 hr long. And yet we found little evidence for long-term efficacy. These results speak against the viability of implicit-bias-oriented diversity trainings as they are currently practiced in real-world settings.

Our study of diversity training on U.S. police officers converges with the mixed findings of diversity trainings in other professional sectors, such as companies or academia (Bezrukova et al., 2016; Carnes et al., 2015; Chang et al., 2019; Devine et al., 2017). These studies indicate that the current generation of diversity trainings are effective at changing minds but less consistent at changing behavior. Our findings also converge with a recent non-peer-reviewed study of an implicit-bias-oriented diversity training of New York City police officers that focused more on behavioral nudges rather than on seeing people as unique individuals (Worden et al., 2020). Like our study, Worden et al.'s study found that training changed minds but not behavior.

We also investigated whether individual differences in officers or educators were linked to training impact. Although officers varied greatly at baseline, the training changed officers a similar amount regardless of background. These findings speak against the possibility of backlash against the training by subgroups of officers. They also converge with meta-analytic null effects of trainee demographics on diversity training outcomes (Bezrukova et al., 2016). Most analyses of the effects of differences between educators also found null effects, which may reflect the high level of standardization in the curriculum. Given that the educators' reference manual was more than 80 pages, there was limited opportunity for educators to express individual differences. The few significant effects we found were observed only in in-person trainings, where individual differences may be expressed more readily. We found that sessions with male–male educator pairs had worse outcomes than educator pairs with at least one woman. Perhaps pairs of male educators advocated for the lessons of the training more weakly because they had less standing to push back against the "cult of masculinity" that defines police culture and reinforces aggression and rigid in-group/out-group distinctions (Brown, 2007; Silvestri, 2017; Smith & Gray, 1985). We also found that sessions with the single ex-law-enforcement educator had better outcomes, although we hesitate to generalize too much from a single individual.

Improving the efficacy of diversity trainings

Diversity trainings show greater efficacy if they are better integrated into a broader organizational strategy and sustained through reinforcement (Bezrukova et al., 2016; Devine & Ash, 2022). In our study, the training was implemented by an external organization and administered separately from the rest of the police departments' activities. The lessons of diversity trainings should be embedded with other organizational initiatives, reinforced by police managers, and evaluated as a part of job performance.

The lack of training efficacy may also be a dosage issue that is characteristic of a lack of organizational commitment. The trainings were delivered in a single sequence without follow-up practice, so there was no

continual reinforcement to ingrain the bias-mitigating strategies into routine habits (W. Wood, 2019). Longer educational experiences are more effective at creating behavioral change (Bezrukova et al., 2016; Bonilla et al., 2021). In U.S. police training, social skills are often deprioritized. Recruits at police academies spend an average of 71 hr on firearms skills and 60 hr on self-defense (Reaves & Trotter, 2017). In contrast, they spend an average of only 9 hr on conflict resolution and 12 hr on cultural diversity skills. That is not that much more time than the total time spent in the Managing Bias training. The limited amount of time spent on diversity training is not endemic to policing; diversity trainings in organizations are often only a couple hours at most. To undo a lifetime of racialized socialization, organizations should invest in more than a day's worth of diversity education.

The lack of training efficacy may reflect an error in teaching approach. The Managing Bias training educated officers on strategies that had been shown to reduce bias in the lab, such as perspective-taking. However, there is a large conceptual gap between compelling individuals to actively take the perspective of another person in the lab and merely instructing people to go out in the world and take the perspective of others. Instructing people to take others' perspectives may have been unpersuasive or unmotivating compared with intensive lab experiences (e.g., Galinsky & Moskowitz, 2000; Herrera et al., 2018). Perhaps alternative approaches to educating about diversity will be more effective in classroom-style instruction. For example, a police training program that educates officers about historical racism and procedural justice has shown greater evidence for efficacy (G. Wood et al., 2020).

Limitations

In this study, we relied on self-reported behavior because it was not logistically possible to obtain other behavioral measures. Self-reports are useful for understanding peoples' personal theories about the actions they have taken. However, those theories may not line up with actual behavior (Kormos & Gifford, 2014; Manfredo & Shelby, 1988). The divergence can be attributed to failures to realize when one has engaged in a certain behavior, to forgetting, or to survey-related measurement issues, such as social desirability or question format. Future studies should complement self-reported behavior with other behavioral measures, such as body camera footage, behavior in simulations, or administrative data on law-enforcement actions (e.g., stops, use of force).

Our study's data collection was also impacted by the COVID-19 pandemic. We had high power to detect changes immediately after training, but the effective samples from the second cohort's 1-month follow-up survey were much smaller. The second cohort's training was also conducted online instead of in person. Although this training program appeared to be quite robust to modality because of a lack of overall differences in the immediate effects of online versus in-person training, we cannot rule out the possibility that the online training was less effective at creating long-term change than an in-person training would be. Future research should examine long-term change using larger samples.

Finally, our study could not describe how characteristics of police departments may be linked to training efficacy. Because of confidentiality concerns, we were not given permission to disclose characteristics of our law-enforcement partners beyond what is already presented in this article. This limitation puts bounds on the generalizability of the data. Future research should document how characteristics of police departments relate to training efficacy.

Conclusion

These findings suggest that single-shot diversity training programs as they are currently practiced in real-world settings may have limited efficacy. Future research may find that greater investment and integration of training programs into organization-wide initiatives could increase their impact. Until stronger evidence is found, practitioners who are interested in addressing bias in organizations should prioritize other reforms for addressing racial disparities in policing.

Transparency

Action Editor: Sylvia Perry
Editor: Patricia J. Bauer
Author Contributions

Calvin K. Lai: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing.

Jaclyn A. Lisnek: Data curation; Formal analysis; Investigation; Methodology; Project administration; Writing – review & editing.

Declaration of Conflicting Interests

C. K. Lai is a consultant and member of the Scientific Advisory Board with Project Implicit, a nonprofit organization and international collaborative of researchers who are interested in implicit social cognition. The authors declared that there were no other potential conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was supported by grants from the Anti-Defamation League and the Russell Sage Foundation.

ORCID iDs

Calvin K. Lai https://orcid.org/0000-0003-2437-9783
Jaclyn A. Lisnek https://orcid.org/0000-0001-8291-8832

Acknowledgments

We are incredibly grateful to the many undergraduate research assistants who helped process the data. We also thank the Anti-Defamation League and our police department organizational partners for their transparency and support in administering the trainings and data collection. We especially thank the many facilitators at the Anti-Defamation League, A. Y., J. G., G. D., D. M., and M. S.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/09567976221150617

Note

1. These strategies were focused on regulating biased behavior rather than reducing implicit bias. Interventions that focus on reducing implicit bias specifically tend to have limited durability and uncertain impacts on behavior (Forscher et al., 2019; Lai et al., 2016).

References

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

Bonilla, S., Dee, T. S., & Penner, E. K. (2021). Ethnic studies increases longer-run academic engagement and attainment. *Proceedings of the National Academy of Sciences, USA*, *118*(37), Article e2026386118. https://doi.org/10.1073/pnas.2026386118

Brown, J. (2007). From cult of masculinity to smart macho: Gender perspectives on police occupational culture. *Sociology of Crime Law and Deviance*, *8*, 205–226. https://doi.org/10.1016/S1521-6136(07)08008-6

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*, *90*(2), 221–230. https://doi.org/10.1097/ACM.0000000000000552

Carr, P. B., Dweck, C. S., & Pauker, K. (2012). "Prejudiced" behavior without prejudice? Beliefs about the malleability of prejudice affect interracial interactions. *Journal of Personality and Social Psychology*, *103*(3), 452–471. https://doi.org/10.1037/a0028849

CBS News. (2019, August 7). *We asked 155 police departments about their racial bias training. Here's what they told us.* https://www.cbsnews.com/news/racial-bias-training-de-escalation-training-policing-in-america/

Chang, E. H., Milkman, K. L., Gromet, D. M., Rebele, R. W., Massey, C., Duckworth, A. L., & Grant, A. M. (2019). The mixed effects of online diversity training. *Proceedings of the National Academy of Sciences, USA*, *116*(16), 7778–7783. https://doi.org/10.1073/pnas.1816076116

Devine, P. G., & Ash, T. L. (2022). Diversity training goals, limitations, and promise: A review of the multidisciplinary literature. *Annual Review of Psychology*, *73*, 403–429. https://doi.org/10.1146/annurev-psych-060221-122215

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*(6), 1267–1278. https://doi.org/10.1016/j.jesp.2012.06.003

Devine, P. G., Forscher, P. S., Cox, W. T. L., Kaatz, A., Sheridan, J., & Carnes, M. (2017). A gender bias habit-breaking intervention led to increased hiring of female faculty in STEMM departments. *Journal of Experimental Social Psychology*, *73*, 211–215. https://doi.org/10.1016/j.jesp.2017.07.002

Dobbin, F., Schrage, D., & Kalev, A. (2015). Rage against the iron cage: The varied effects of bureaucratic personnel reforms on diversity. *American Sociological Review*, *80*(5), 1014–1044.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, *117*(3), 522–559. https://doi.org/10.1037/pspa0000160

Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*, *72*, 133–146.

Frey, E., & Rogers, T. (2014). Persistence: How treatment effects persist after interventions stop. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 172–179. https://doi.org/10.1177/2372732214550405

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724.

Glaser, J. (2015). *Suspect race: Causes and consequences of racial profiling*. Oxford University Press.

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE*, *13*(10), Article e0204494. https://doi.org/10.1371/journal.pone.0204494

James, T. (2017, December 23). Can cops unlearn their unconscious biases? *The Atlantic*. https://www.theatlantic.com/politics/archive/2017/12/implicit-bias-training-salt-lake/548996/

Kormos, C., & Gifford, R. (2014). The validity of self-report measures of proenvironmental behavior: A meta-analytic review. *Journal of Environmental Psychology*, *40*, 359–371. https://doi.org/10.1016/j.jenvp.2014.09.003

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., Blanchar, 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

Lueke, A., & Gibson, B. (2015). Mindfulness meditation reduces implicit age and race bias: The role of reduced automaticity of responding. *Social Psychological and Personality Science*, *6*(3), 284–291. https://doi.org/10.1177/1948550614559651

Ma, D. S., Correll, J., Wittenbrink, B., Bar-Anan, Y., Sriram, N., & Nosek, B. A. (2013). When fatigue turns deadly: The association between fatigue and racial bias in the decision to shoot. *Basic and Applied Social Psychology*, *35*(6), 515–524. https://doi.org/10.1080/01973533.2013.840630

Manfredo, M. J., & Shelby, B. (1988). The effect of using self-report measures in tests of attitude—Behavior relationships. *The Journal of Social Psychology*, *128*(6), 731–743. https://doi.org/10.1080/00224545.1988.9924553

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology*, *65*(3), 469–485. https://doi.org/10.1037/0022-3514.65.3.469

Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*(1), 533–560. https://doi.org/10.1146/annurev-psych-071620-030619

Perry, S. P., Murphy, M. C., & Dovidio, J. F. (2015). Modern prejudice: Subtle, but unconscious? The role of bias awareness in Whites' perceptions of personal and others' biases. *Journal of Experimental Social Psychology*, *61*, 64–78. https://doi.org/10.1016/j.jesp.2015.06.007

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

President's Task Force on 21st Century Policing. (2015, May). *Final report of the President's Task Force on 21st Century Policing*. Office of Community Oriented Policing Services.

Rattan, A., & Georgeac, O. A. M. (2017). Understanding intergroup relations through the lens of implicit theories (mindsets) of malleability. *Social and Personality Psychology Compass*, *11*(4), Article e12305. https://doi.org/10.1111/spc3.12305

Reaves, B. A., & Trotter, A. L. (2017). The state of law enforcement training academies. *Sheriff & Deputy*, *69*(1), 44–47.

Rubinstein, R. S., Jussim, L., & Stevens, S. T. (2018). Reliance on individuating information and stereotypes in implicit and explicit person perception. *Journal of Experimental Social Psychology*, *75*, 54–70. https://doi.org/10.1016/j.jesp.2017.11.009

Sierra-Arévalo, M. (2021). American policing and the danger imperative. *Law & Society Review*, *55*(1), 70–103. https://doi.org/10.1111/lasr.12526

Silvestri, M. (2017). Police culture and gender: Revisiting the 'cult of masculinity.' *Policing: A Journal of Policy and Practice*, *11*(3), 289–300. https://doi.org/10.1093/police/paw052

Smith, D. J., & Gray, J. (1985). *Police and people in London*. Gower Publishing.

Soss, J., & Weaver, V. (2017). Police are our government: Politics, political science, and the policing of race–class subjugated communities. *Annual Review of Political Science*, *20*(1), 565–591. https://doi.org/10.1146/annurev-polisci-060415-093825

Spencer, K. B., Charbonneau, A. K., & Glaser, J. (2016). Implicit bias and policing. *Social and Personality Psychology Compass*, *10*(1), 50–63. https://doi.org/10.1111/spc3.12210

Stoughton, S. (2014). Law enforcement's "warrior" problem. *Harvard Law Review Forum*, *128*, Article 225. https://harvardlawreview.org/2015/04/law-enforcements-warrior-problem/

Wood, G., Tyler, T. R., & Papachristos, A. V. (2020). Procedural justice training reduces police use of force and complaints against officers. *Proceedings of the National Academy of Sciences, USA*, *117*(18), 9815–9821. https://doi.org/10.1073/pnas.1920671117

Wood, W. (2019). *Good habits, bad habits: The science of making positive changes that stick*. Macmillan.

Worden, R. E., McLean, S. J., Engel, R. S., Cochran, H., Corsaro, N., Reynolds, D., Najdowski, C. J., & Isaza, G. T. (2020). *The impacts of implicit bias awareness training in the NYPD*. https://www1.nyc.gov/assets/nypd/downloads/pdf/analysis_and_planning/impacts-of-implicit-bias-awareness-training-in-%20the-nypd.pdf

Yurek, L. A., Vasey, J., & Sullivan Havens, D. (2008). The use of self-generated identification codes in longitudinal research. *Evaluation Review*, *32*(5), 435–452. https://doi.org/10.1177/0193841X08316676