

Developing & delivering effective anti-bias training: Challenges & recommendations

Evelyn R. Carter, Ivuoma N. Onyeador, & Neil A. Lewis, Jr.

abstract

Organizations invest nearly \$8 billion annually in diversity training, but questions have arisen about whether training actually reduces biased attitudes, changes behavior, and increases diversity. In this article, we review the relevant evidence, noting that training should be explicitly aimed at increasing awareness of and concern about bias while at the same time providing strategies that attendees can use to change their behavior. After outlining five challenges to developing and delivering training that meets these goals, we provide evidence-based recommendations that organizations and facilitators can use as a blueprint for creating anti-bias training programs that work. One recommendation is to couple investment in anti-bias training with other diversity and inclusion initiatives to help ensure that the billions spent each year yield meaningful change.

What do Starbucks, Delta Airlines, and the Napa Valley Wine Train have in common? Over the past five years, in response to public outcries over racist incidents between employees and customers, all three have invested considerable time and money in diversity training programs.¹⁻³ They are not alone: it has been estimated that organizations invest nearly \$8 billion each year in such training.⁴

Core Findings

What is the issue?

Organizational diversity initiatives tend to include explicit anti-bias training in the hope of alleviating harmful biases among employees in the workforce. But is this money well spent? Recent research highlights the need to redevelop anti-bias initiatives to do more than simply create awareness of biases or prompt defensive reactions. Instead, anti-bias training should increase awareness of bias and its lasting impact, plant seeds that inspire sustained learning, and teach skills that enable attendees to manage their biases and change their behavior.

How can you act?

Selected recommendations include:

- 1) Including training as part of a broader diversity and inclusion strategy
- 2) Prioritizing the learning needs of the primary audience over positive reviews

Who should take the lead?

Researchers and organizational leaders

The training can take different forms, but all programs implicitly or explicitly aim to reduce people's biased attitudes and behaviors during everyday activities in organizations, such as when interacting with customers or colleagues or making hiring and promotion decisions. In this article, the term *bias* refers to differences in how people are viewed and treated as a result not of their individual characteristics but of group features, such as gender, skin color, or ethnicity. Bias can be positive for some people (conferring preference or granting privilege) or negative (being discriminatory or otherwise causing disadvantage) and can influence interactions between individuals or be expressed through policies and practices that affect one group differently from another.⁵

Organizations have varied reasons for pouring money into diversity training. After all, the well-documented consequences of bias in companies are far-reaching. Bias can harm the mental and physical health of employees who experience it,⁶⁻⁸ interfere with their performance and engagement,^{9,10} and undermine their professional development and promotion.^{11,12} Bias also undercuts efforts to increase inclusion and diversity in who gets hired and fills management positions. Conversely, employees view companies that explicitly commit to recognizing and celebrating diversity as more trustworthy.¹³ Further, what is often called the "business case" for diversity holds that reducing bias and increasing diversity has the potential to increase profitability for companies, improve their reputation, and limit internal and external liability.^{14,15}

The problem is that research on how well diversity training succeeds in reducing bias and

increasing diversity has produced mixed results, leading many investigators to conclude that it does not work.^{16,17} It seems likely, however, that the inconsistent results stem from differences in how the training is carried out. Some programs, for instance, focus mainly on making people aware of their own biases or on providing attendees with a laundry list of things they should not do (often with little explanation as to why their behavior should change). Others design situations so that attendees can experience firsthand what it is like to be a target of bias. Some training sessions are led by experts in the subjects of diversity and inclusion, whereas others are led by facilitators with little expertise or experience in the domain. Finally, whereas some training programs are bolstered by resounding institutional support, others are perceived as a specialized interest of a small corner of the organization, garnering less support from leadership and participation by employees as a result. These differences could affect outcomes and result in mixed findings as to the conditions that make training succeed or fail. Indeed, only some of the above strategies have been empirically shown to reduce bias.

Thus, we argue that diversity training should go beyond telling people that bias exists or creating uncomfortable experiences that are more likely to prompt defensiveness than learning. Rather, the most effective training is anti-bias training that is designed to increase awareness of bias and its lasting impact, plant seeds that inspire sustained learning, and teach skills that enable attendees to manage their biases and change their behavior. Although a dearth of consistent evidence of success has led many to conclude that training does not work,^{16,17} empirical research exists that can provide a blueprint for how to build a training program that does.

In this article, we explain the logic behind our argument, present five challenges to developing and delivering effective anti-bias training, and offer evidence-based recommendations for how to overcome those challenges (see Table 1 for a summary). We acknowledge, however, that it is not easy to develop a single training program that will reduce the bias that affects all of the myriad groups in an organization.

Table 1. Summary of challenges, recommendations, & implementation agents

Challenge	Recommendation	Implementation Agent
Being realistic about what training can accomplish on its own	Develop a comprehensive diversity and inclusion strategy with training as one component. Determine which goals require alternative strategies.	Organizational executives responsible for strategic planning initiatives
Selecting the proper goals for the programs	Tailor training to match the desired outcomes for individuals, groups, and the organization. Focus on increasing awareness of bias and changing behavior.	Organizational executives who choose the goal; facilitator who tailors training accordingly
Deciding how to manage attendee discomfort	Do not let discomfort dissuade the organization from pursuing training; discomfort is an important part of learning how to communicate across differences.	Training facilitator
Minimizing counterproductive effects of discussing bias	Teach attendees concrete strategies for managing their bias, but do not overdo it: emphasize just two or three.	Training facilitator
Demonstrating impact	Develop and execute a plan for evaluating the efficacy of training.	Organizational executives and training facilitator

Because bias and stereotypes can take many forms,¹⁸ a training program that helps to reduce bias against one group may not necessarily be effective to the same degree for another. For instance, whereas diversity efforts that highlight group differences improve conditions for groups present in moderate numbers, such as White women, efforts that emphasize equal treatment and minimize the salience of group differences improve conditions for groups with low representation, such as Black people.¹⁹ We also note the need for more research to understand whether a training program built using our recommendations would reduce bias against understudied groups, such as people who are disabled or obese.

Challenge 1: Being Realistic About What Training Can Accomplish on Its Own

Because of budgetary and time considerations, diversity training is often offered as a one-time opportunity. Yet a company that relies solely on a single training session to combat bias is doomed to fail. One major reason is that bias is multifaceted, born of a combination of an individual's exposure to stereotypes about and direct or indirect experiences with people from different groups. This complexity makes bias difficult to eradicate. Indeed, an investigation

of 17 different bias-reduction interventions found that only eight reduced participants' implicit preference for White people over Black people.²⁰ Further, the effects of even the most effective interventions (such as exposing participants to people from another group who behave counter to stereotypes and providing people with strategies they can use to mitigate bias) had worn off just 24 hours later.²¹ This analysis and other research suggest that without consistent reinforcement, trainees' biases will rebound after people return to the environments that reinforce those biases.²²

For example, after attending an anti-bias training session, a hiring manager may be excited about implementing new bias-management strategies. Yet this manager must balance the goal of recruiting and offering jobs to a diverse group of candidates with the competing need to fill open positions swiftly. If the company has not invested in targeted recruiting of candidates from historically underrepresented groups or has not reduced the urgency of filling a particular position, the newfound knowledge and motivation to counteract bias will not be enough to combat the pressing need to fill the position and stiffen the manager's resolve to hold out for a qualified person who also adds to the company's diversity.²³ Indeed, research suggests that accountability structures produce better results

“organizations cannot rely on employees’ goodwill alone to change the culture”

than anti-bias training by itself.²⁴ (Accountability structures consist of programs, staff positions, or groups that have explicit responsibility for meeting such goals as increasing the representation of minorities.) In other words, although anti-bias training can be a great way to raise awareness of bias and provide people with strategies to fight it, training alone cannot guarantee change once attendees leave a session.

Recommendation: Include Training as Part of a Broader Diversity & Inclusion Strategy

Organizations must ensure that training is accompanied by investments in structural changes—alterations in organizational policies and ways of operating—that will help sustain learning and facilitate behavioral change.^{25,26} Why are structural changes imperative even within an organization that is full of employees personally motivated to be egalitarian? One reason is that prejudice is facilitated by the daily consumption of information that confirms and reifies stereotypes.^{27,28} Another is that, as is illustrated by the hiring manager example above, many factors compete with an individual’s motivation to prioritize diversity, manage bias, and promote inclusion. Combating prejudice requires similarly multifaceted efforts: organizations cannot rely on employees’ goodwill alone to change the culture.^{29,30}

A meta-analysis of 260 studies of anti-bias training highlights the importance of going beyond having a training program on its own. It found that training was more effective when it was integrated into a broader strategy ($g = 0.57$; see note A for more information on the statistics) than when it was provided as a stand-alone program ($g = 0.36$).³¹ An integrated strategy might introduce anti-bias training along with systems that hold individuals responsible for reducing bias, improve processes for responding to bias incidents, and provide networking opportunities for employees from underrepresented groups. Training is more effective in this broader context in part because the integrated strategy addresses the variety of ways that bias

can undermine an organization’s diversity and inclusion efforts. Moreover, organizations that use a comprehensive strategy communicate the message that bias is malleable and can be changed over time, which can encourage individuals to adopt bias-management strategies.^{32,33}

Organizations with the goal of integrating training into their overall diversity and inclusion strategy could first collect data to understand the representation and dispersion of people from different groups throughout the organization, employees’ perceptions of inclusion, and where diversity-related failures in processes might be occurring (such as during the hiring process or during performance reviews). The data would help to identify which training topics would have the greatest impact on the organization’s employees. The organization could then create a plan in which training for employees and leaders is one component, alongside the establishment of a committee responsible for overseeing and reporting on progress. After training, organizations could provide suggestions for additional reading to attendees to further their learning and establish behavioral nudges before key decisions to remind attendees of the strategies to which they committed.³⁴ Such nudges could include reminding managers to avoid giving personality feedback before performance reviews are written or asking recruiters to reflect on key job requirements before they discuss their assessments of job candidates. In these ways, organizations can ensure the training lessons extend beyond the session itself and influence employees’ everyday behavior.

Challenge 2: Selecting the Proper Goals for the Programs

Researchers—such as Frank Dobbin and his colleagues—have analyzed a combination of laboratory experiments and organizational data and concluded that diversity training is ineffective.³⁵ For example, Dobbin and his coauthors

examined how six different diversity-related initiatives (training, audits, networking programs, mentoring programs, a task force, and a designated diversity officer) affected manager-level diversity in more than 800 companies. Mentoring programs and initiatives that provided organizational oversight for diversity through an individual or a task force increased management diversity by nearly 40%. In contrast, the impact of training was generally negligible and for some groups was counterproductive, contributing to a 5% decrease in representation for Black women and a 5%–10% increase for Black men and Hispanic women. The authors concluded that training is not effective for increasing manager diversity. However, an alternative conclusion is not necessarily that anti-bias training is ineffective but rather that it is not a panacea.

Many factors influence employee diversity, promotion rates, and other outcomes. Consider the goal of increasing manager-level diversity. If White employees have more opportunities to lead high-visibility projects and have greater access to key decisionmakers than employees of color do, this imbalance will necessarily affect the pipeline of qualified employees who can become managers.³⁶ Training cannot fix these disparities. Similarly, anti-bias training cannot promise comprehensive change that will eliminate all bias that occurs during performance reviews³⁷ or all overt or subtle biases in other interactions.^{38–40} Instead, it is best positioned to achieve more modest and immediate outcomes, such as educating people about how biases can manifest or motivating people to change their behavior.

Recommendation: Focus Training on Both Increasing Awareness of Bias & Providing Strategies for Changing Behavior

Some researchers suggest that the closest analog of training is teaching and that organizations considering diversity training should first identify the learning outcomes of interest.⁴¹ In the meta-analysis mentioned above, Katerina Bezrukova and colleagues identified four potential kinds of outcomes from diversity training: attendee reactions (self-reported feelings toward the facilitator or training), cognitive learning (knowledge gained about the topic,

such as where bias comes from and how it contributes to inequity in society), behavioral learning (skill development, as judged by self-reports, observations of managers or trainers, or responses to hypothetical scenarios), and attitudinal/affective learning (self-reported attitudes toward stigmatized group members and beliefs in one's own ability to become less biased).³¹ Overall, training had the largest immediate impact on generating positive attendee reactions toward the facilitator and the training itself ($g = 0.61$). It also increased, to varying degrees, awareness of bias (cognitive learning; $g = 0.57$), skill at behaving in less-biased ways (behavioral learning; $g = 0.48$), and positive feelings toward groups that were experiencing discrimination (attitudinal/affective learning; $g = 0.30$). Over time, however, the effects decayed for all outcomes except awareness of bias (cognitive learning). Thus, one outcome that anti-bias training can reliably affect in the long term is what attendees learn and retain about bias.

If awareness of bias is the main outcome that persists, should organizations even bother trying to change behavior as well? Based on the data, the short answer is yes. In a separate analysis, Bezrukova and her colleagues asked what works better for changing attitudes and behavior: diversity training that aims to increase attendees' awareness of their biases and cultural assumptions (awareness-based training), programs that help attendees learn to monitor and change their behavior (behavior-based training), or a combination of the two?³¹ They found that focusing on awareness is useful but should not be the only focus of training. Awareness-based training produced the smallest changes in attitudes and behavior overall ($gs = 0.22$ and 0.35 , respectively), whereas behavior-based training was significantly more effective at changing both attitudes ($g = 0.41$) and behavior ($g = 0.53$). The training programs that incorporated both awareness-based and behavior-based elements were about as effective as behavior-based training at changing attitudes ($g = 0.40$) and behavior ($g = 0.54$). In summary, anti-bias training is least effective when it focuses only on raising awareness of bias: the best strategy is either to focus on teaching attendees strategies for changing biased behavior or to do that

and also incorporate elements that will raise people's awareness of their bias and the effects it can have.⁴² We favor the combined approach because it can help people to understand why they should want to change their behavior.

Challenge 3: Deciding How to Manage Attendee Discomfort

A critical component of anti-bias training is discussing disparities between groups. These conversations can be uncomfortable. For example, people from racial minority and majority groups both find discussions about race to be challenging and thus may avoid intergroup conversations about race to prevent the accompanying anxiety and discomfort.^{43–46} Although avoidance certainly eases strain between people in the short term, avoidance is not always possible during anti-bias training and, more generally, when working in a diverse organization. During training, some attendees may become defensive and belittle the training content or even try to undermine the facilitator in an effort to deflect negative emotions.^{47,48} Facilitators must be prepared to contend with the various reactions that emerge during training.

Facilitators must also navigate the tricky dynamics related to differences in how people from various groups perceive the prevalence of bias and in their willingness to engage in the training activities. White participants, for instance, may believe that racism is less prevalent in modern society than Black participants do.^{49,50} Also, recognizing one's own bias is distressing, particularly for majority-group members who do not want to appear prejudiced.^{51,52} Some attendees may be reluctant to acknowledge and accept that they have biases that affect their behavior, or they may be actively hostile to the very notion of anti-bias training.^{53,54} Other attendees may generally agree that there is a need to improve organizational inclusion and reduce bias.

These divergent perceptions and attitudes can, in turn, produce different expectations for what anti-bias training should cover. Some attendees may prefer to have an introductory conversation

to provide them with foundational knowledge about the prevalence of bias before diving into strategies for mitigating it. Other attendees may be eager for more complex discussions, looking to anti-bias training as an opportunity to validate their experiences with bias and to discuss meaningful changes their organization can enact. These dynamics make it challenging to design a training program that all attendees will react to positively. It is reasonable and understandable that organizational leaders want employees to have good reactions to training programs, but this goal should not be pursued at the expense of other learning outcomes.

Recommendation: Prioritize the Learning Needs of the Primary Audience, Not Positive Reviews

As part of establishing the desired learning outcomes for anti-bias training, organizations must identify the primary audience for any particular anti-bias training session. Potential targets of bias will undoubtedly be in the audience of any training session, and their needs may be very different from those of attendees who are potential perpetrators of bias. For example, people of color learn about racism and talk about racism at a younger age than White individuals do.^{55,56} Thus, a discussion of the existence of bias and the importance of managing it may fall flat for people of color, who instead may gain more from learning about how to cope with the bias they face or from discussing organizational procedures for reporting discrimination. Although one training session cannot address all these topics effectively, a scaffolded approach to learning—through a series of workshops, comprehensive diversity and inclusion efforts (as described in Challenge 1), or both—can ensure that everyone within an organization gets the education and support they need.

After the identity of the primary audience has been established, the next question to consider is the potential reactions that attendees will have to training. Some facilitators will prioritize designing training programs that attendees will evaluate positively over training programs that achieve learning outcomes. Several studies' findings suggest, however, that this impulse is misguided. Certainly, the perception of too



\$8 billion is spent on diversity training annually



50% of US organizations are expected to offer implicit bias training in the future

“a moderate amount of discomfort is a critical catalyst for the introspection that can guide a person toward more egalitarian behavior in the future”

much threat can activate defensiveness that stifles learning and growth.⁵⁷ Yet research has shown that a moderate amount of discomfort is a critical catalyst for the introspection that can guide a person toward more egalitarian behavior in the future.⁵⁸ This research aligns with other evidence that showing people how bias has influenced their decisions can reduce subsequent discrimination⁵⁹ and increase sensitivity to subtle forms of discrimination.⁶⁰ When facilitators and organizations make winning audience approval a key outcome, they risk neglecting content that may produce long-lasting learning (that is, ongoing sensitivity to the occurrence of bias) and behavioral change.

Another reason to prioritize attendees' learning needs over their immediate reactions is that attendees' positive or negative perceptions of training do not necessarily relate to whether they learn the concepts presented during the experience.⁶¹ For instance, attendees react more positively to training that includes a variety of components (such as interactive discussion, video, and lecture-based content; $g = 0.73$) than to training that relies on one component ($g = 0.59$).³¹ Despite this preference, training programs with many components, as compared with those with one component, do not produce statistically significant differences in behavioral learning (many: $g = 0.51$; one: $g = 0.39$), cognitive learning (many: $g = 0.54$; one: $g = 0.67$), or attitudinal/affective learning (many: $g = 0.30$; one: $g = 0.27$). Thus, although audiences may prefer training programs with multiple components to training programs with one component, the multifaceted approach does not yield meaningful change on any learning outcomes.

Finally, organizations often make training voluntary rather than mandatory to generate positive audience reaction. In some ways, this strategy is effective: overall, people respond more favorably to voluntary versus mandatory training

($gs = 0.71$ and 0.37 , respectively). However, this approach undermines other outcomes that may be important for long-term change, such as behavioral learning.³¹ Specifically, when training is voluntary, behavioral learning is significantly lower compared with when training is mandatory ($gs = 0.42$ and 0.63 , respectively), perhaps because those who could benefit most from the training avoid attendance.^{62,63} Moreover, making training mandatory is a simple way to demonstrate that anti-bias training and, more broadly, diversity and inclusion efforts are important to the organization.

Challenge 4: Minimizing Counterproductive Effects of Discussing Bias

A major goal of anti-bias training is to increase awareness of the ways that bias manifests in society, organizations, and individuals. However, several scholars have identified unexpected, often ironic effects that sometimes result from attempts to teach about and address bias. For example, people more readily dismiss claims of workplace discrimination in organizations that explicitly value diversity compared with claims in organizations that do not.⁶⁴ Communicating that the solution to bias is simple and can be addressed by just offering a training course can decrease empathy for victims of bias.⁶⁵ Additionally, referring to bias as ubiquitous may, ironically, produce the perception that bias is acceptable (for instance, “It must be OK if everyone has it”)⁶⁶ or that bias-reduction efforts are futile (for instance, “You can’t succeed because bias is too widespread to root out”).

Beyond these challenges, the findings of recent research complicate notions of how to talk about implicit bias (sometimes called *unconscious bias*) in particular. An estimated 20% of U.S. organizations offer training meant to combat implicit bias, and the number is expected to increase to 50% in the near future.⁴

Although this training can be effective,⁶⁷ the way that implicit bias is framed can significantly affect people's motivation to challenge the discrimination that results from it. In a series of studies, participants read a definition that framed racial bias as "implicit and unconscious" or "explicit and conscious" before reading about a discriminatory incident. Those who read about implicit racial bias perceived less intent in the incident, and they therefore perceived the incident as less harmful and the perpetrator as less worthy of blame and punishment.^{68,69} Because the very framing of anti-bias training content can undermine the critical understanding of bias necessary for motivating behavioral change, facilitators must pay close attention to their framing.

Recommendation: Help Attendees Create a Plan for Behavioral Change

When building anti-bias training with the goal of circumventing the ironic consequences of talking about bias, facilitators would do well to consider classic social psychological research on persuasion. Persuasion is often most successful when people are presented with a moderately disturbing outcome and strategies they can use to avoid that outcome.⁷⁰ Indeed, anti-bias training is a persuasive endeavor, designed to present attendees with information that will motivate them to change their attitudes and behavior. Facilitators must go beyond relaying information about what bias is and propose concrete bias-management strategies so that attendees believe they have the ability to enact the new behaviors.^{71–73}

Although myriad strategies exist for managing and reducing bias, facilitators should be careful not to overwhelm attendees with too many of them. General research on interventions has compared the effect of different numbers of recommendations on behavioral change.⁷⁴ An interesting finding was that interventions that provided one recommendation had a low impact on behavior ($d = 0.12$), but so did interventions that provided four or more recommendations ($d = 0.14$). (See note A for a discussion of d statistics.) Giving just one recommendation suggested the problem was not important, but four or more overwhelmed participants. Instead,

two to three recommendations emerged as the ideal number for prompting change ($d = 0.27$), particularly for participants who were less motivated to alter their ways. Throughout a training program, facilitators may suggest a broad range of strategies to manage and reduce bias; however, by the end, facilitators who want to ensure that their attendees change their behaviors should encourage them to focus on two or three behaviors.

Challenge 5: Demonstrating Impact

Clearly, it is critical that facilitators put substantive thought into the framing and content of anti-bias training. However, this effort is for naught if the impact of training is not evaluated. Yet barriers can interfere with such evaluation. For instance, although collecting data on the efficacy of training is a crucial step in assessment, accessing organizational data, which could show the impact that training has on key outcomes, can be challenging. In addition, if participation in postworkshop surveys is low, the paucity of data can undermine the ability to draw statistically informative conclusions, and a focus on unrealistic outcomes (as discussed in Challenge 2) can give the misleading impression that nothing was achieved.

Unfortunately, a lack of data showing that anti-bias training is effective can lead people to perceive anti-bias training as ineffective. Perhaps worse, without data demonstrating a return on an organization's investment in anti-bias training, organizational leaders will be skeptical when they receive future requests to fund diversity and inclusion initiatives. Without data to point to, facilitators—and the organizational representatives who hire them—will be ill-equipped to make a strong case for why the training is necessary and worthwhile.

Recommendation: Measure Efficacy

Any plan for delivering training should, from the outset, also include a plan for assessing how well the program's goals are met.⁷⁵ The program goals, of course, should be established at the beginning of the collaboration between the organization and the facilitator to

ensure that the objectives of both align and that the content included in the training serves the selected goals. When deciding on the outcomes of interest, organizations can also consider what their comprehensive diversity and inclusion strategy should be (as discussed in Challenge 1).

Case studies from a variety of industries provide insight into what anti-bias training can potentially accomplish,^{67,76} and empirical evaluations conducted once training was complete have bolstered these findings. For example, The Ohio State University College of Medicine identified a goal of increasing the diversity of its incoming class of medical students. The college developed a plan to have admissions committee members complete anti-bias training; the plan included provisions for data collection at the end so that the effectiveness of the training could be evaluated. Per the plan, the college's admissions committee members took an Implicit Association Test and attended a training session about how implicit bias affects decisionmaking.⁷⁷ During the next admission cycle, the college saw a 26% increase in the number of underrepresented minority students who matriculated to the institution, compared with the admissions cycle from the previous year. This difference was not statistically significant; nonetheless, qualitative responses from an evaluation completed by committee members gave some insight into the way the training affected them. Specifically, they described being more cognizant of how bias influenced their reactions to candidates during the interview process. By identifying the pretraining goal of increasing the diversity of the incoming class of medical students, the college was able to develop a training approach that positively shaped the behavior of the admissions committee and that also included a plan to collect the data needed to assess outcomes in the months that followed.

Discussion

Social scientists have amassed extensive knowledge about the features that are important for effective anti-bias training. However, the proliferation of training and the lack of standardization in its delivery undermine the ability to measure the efficacy of training across organizations and

"The field of anti-bias training presents a ripe opportunity for collaboration between researchers and practitioners"

over time. Facilitators and organizations who want to improve outcomes should follow the evidence-based recommendations presented in this article for raising awareness of bias and sparking a desire to behave in ways that manage it. Nevertheless, anti-bias training cannot and should not be viewed as a singular opportunity to educate and change people. Rather, it must be integrated into efforts to meet a broader commitment to improving diversity and inclusion. Finally, every organization implementing these recommendations should measure the impact of what it has tried and adjust accordingly, taking into account the specific context, audience, and desired outcomes.

A Call to Continued Action for Researchers

The future success of anti-bias training turns on whether researchers use their skills and experience in program development and evaluation to build strong training programs. This call for researchers to engage in translational research is not new: over a decade ago, Elizabeth Paluck cautioned that "by and large, scholars and practitioners have passed up the opportunity for a collaborative project that could harness this widespread intervention to improve the theory and practice of prejudice reduction and social inclusion."⁷⁸ The field of anti-bias training presents a ripe opportunity for collaboration between researchers and practitioners. We echo Paluck's recommendations for how to achieve this collaboration, which should begin by (a) establishing that anti-bias training can cause attitude and behavioral changes, (b) measuring efficacy in ways that go beyond self-report, and (c) conducting research in a variety of populations and settings to gain a deeper understanding of how to make training most effective. Moreover, researchers must continue to explore the various determinants of bias^{18,79}

and use these insights to identify bias-reduction tactics that are specific to certain groups⁸⁰ as well as tactics that have a broader reach.⁸¹

A Call to Action for Organizations Seeking Anti-Bias Training

People who want to bring anti-bias training to their organization may be overwhelmed by the task of establishing criteria that will identify, from the wealth of facilitators available, the one(s) best suited for their particular group and situation. With that in mind, we recommend that they consider the following factors, regardless of whether they plan to hire an outside facilitator or use internal resources.

The first factor is *organizational readiness*. It is not worth pursuing anti-bias training if the organization lacks a true commitment to increasing diversity and fostering inclusion. Does your organization have a diversity and inclusion strategy, of which training is one component? Conversely, are there people in the organization who will undermine the efficacy of training, such as resistant leaders or employees who are openly antagonistic to diversity and inclusion efforts? If resistance is an issue, company resources may be better spent by first investing in structural changes to policy and practices that can lower the resistance.

Leaders should also carefully consider *facilitator experience*, including subject matter expertise and experience managing the myriad audience reactions to training (described in Challenge 3). Finally, any person who advocates for anti-bias training within an organization must have *a plan for measuring the impact*. Identifying the desired key outcomes for learning and subsequent change will likewise be important when partnering with the facilitator to build the training.

A Call to Action for Anti-Bias-Training Facilitators

Bias can seem abstract, and anti-bias-training facilitators have the potential to make the importance of managing and reducing bias real for attendees. Facilitators have an opportunity to emphasize how biases in our society are

rooted in a long-standing hierarchy and maintained by present-day beliefs and practices.^{82,83} These opportunities to reflect on and discuss bias can influence whether and how members of majority groups notice and make meaning of the experiences of members of stigmatized groups.^{49,84–87} Anti-bias-training facilitators therefore carry a great responsibility, and those interested in entering this field should take this role seriously.

Merely having a passion for diversity and inclusion does not make a person an effective facilitator. Navigating the challenges described in this article, especially those related to the unpredictable and sometimes ironic consequences of discussing bias, requires a deep understanding of the research on bias, diversity, and inclusion as well as experience facilitating training with a variety of audiences. In an early version of this article, a reviewer commented that the challenges and recommendations for facilitators had a Goldilocks feel to them: Do not do too much, do not do too little; present enough information to motivate people but not so much that you overwhelm them. Walking that line is very difficult. And yet it is acceptable, even appropriate, that there be a high bar of entry for facilitating these training programs. No one expects people to practice medicine or law after reading a few articles. Similarly, it does and should require significant time and energy to learn to become an effective anti-bias-training facilitator.

Conclusion

The recommendations in this article should guide the next decade—and beyond—of anti-bias training. As systemic inequities and biases become more entrenched, it is not reasonable to expect a transformation to come from training alone. However, a well-designed training program can be a catalyst that produces ripple effects within an organization, a community, and beyond. Researchers and facilitators should come together to share insights on ways to make anti-bias training as effective as it can be and then use those insights to create a less biased, more equitable world.

endnote

- A. From the editors to nonscientists: For any given data set, the statistical test used depends on the number of data points and the type of measurement being used, such as proportions or means. Hedges's g is a measure of effect size, with interpretation as follows: $g = 0.20$ is considered a small effect, $g = 0.50$ is considered a medium effect, and $g = 0.80$ is considered a large effect. Cohen's d also measures effect size. Typically, $d = 0.2$ is small, $d = 0.5$ is medium, and $d = 0.8$ is large.

author affiliation

Carter: Paradigm Strategy Inc. Onyeador: Yale University and Kellogg School of Management. Lewis: Cornell University and Weill Cornell Medical College. Corresponding author's e-mail: evelyn.carter10@gmail.com.

author note

This research was supported by National Science Foundation Social, Behavioral, and Economic Sciences Postdoctoral Research Fellowship 1809370 to Ivuoma N. Onyeador.

references

- Bhattacharjee, R. (2015, August 25). 'We were 100 percent in the wrong': Napa Wine Train exec apologizes to Black women thrown off train for laughing loudly. Retrieved from <https://www.nbcbayarea.com/news/local/We-Were-100-in-the-Wrong-Napa-Wine-Train-Exec-Apologizes-to-Black-Women-Thrown-Off-Train-for-Laughing-Loudly-322841941.html>
- Shen, L. (2017, January 18). Delta adds diversity training for 23,000 crew members. *Fortune*. Retrieved from <https://fortune.com/2017/01/18/delta-diversity-training/>
- Starbucks. (2018). The third place: Our commitment renewed. Retrieved from <https://stories.starbucks.com/stories/2018/thethirdplace/>
- Lipman, J. (2018, January 25). How diversity training infuriates men and fails women. *Time*. Retrieved from <http://time.com/5118035/diversity-training-infuriates-men-fails-women/>
- Adams, G., Edkins, V., Lacka, D., Pickett, K. M., & Cheryan, S. (2008). Teaching about racism: Pernicious implications of the standard portrayal. *Basic and Applied Social Psychology*, 30, 349–361.
- Adam, E. K., Heissel, J. A., Zeiders, K. H., Richeson, J. A., Ross, E. C., Ehrlich, K. B., . . . Eccles, J. S. (2015). Developmental histories of perceived racial discrimination and diurnal cortisol profiles in adulthood: A 20-year prospective study. *Psychoneuroendocrinology*, 62, 279–291.
- Clark, R., Anderson, N. B., Clark, V. R., & Williams, D. R. (1999). Racism as a stressor for African Americans: A biopsychosocial model. *American Psychologist*, 54, 805–816.
- Jones, K. P., Peddie, C. I., Gilrane, V. L., King, E. B., & Gray, A. L. (2016). Not so subtle: A meta-analytic investigation of the correlates of subtle and overt discrimination. *Journal of Management*, 42, 1588–1613.
- Levy, D. J., Heissel, J. A., Richeson, J. A., & Adam, E. K. (2016). Psychological and biological responses to race-based social stress as pathways to disparities in educational outcomes. *American Psychologist*, 71, 455–473.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.
- Hewlett, S. A. (2014, January 22). Cracking the code that stalls people of color. *Harvard Business Review*. Retrieved from <https://hbr.org/2014/01/cracking-the-code-that-stalls-multicultural-professionals>
- Zenger, J., & Folkman, J. (2019, June 25). Research: Women score higher than men in most leadership skills. *Harvard Business Review*. Retrieved from <https://hbr.org/2019/06/research-women-score-higher-than-men-in-most-leadership-skills>
- Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Dittmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, 94, 615–630.
- Bregman, P. (2012, March 12). Diversity training doesn't work. *Harvard Business Review*. Retrieved from <https://hbr.org/2012/03/diversity-training-doesnt-work>
- Eswaran, V. (2019, April 29). The business case for diversity in the workplace is now overwhelming. Retrieved from <https://www.weforum.org/agenda/2019/04/business-case-for-diversity-in-the-workplace/>
- Belluz, J. (2018, May 29). Companies like Starbucks love anti-bias training. But it doesn't work—and may backfire. Retrieved from <https://www.vox.com/science-and-health/2018/4/19/17251752/philadelphia-starbucks-arrest-racial-bias-training>
- Dobbin, F., & Kalev, A. (2018). Why doesn't diversity training work? The challenge for industry and academia. *Anthropology Now*, 10(2), 48–55.
- Zou, L. X., & Cheryan, S. (2017). Two axes of subordination: A new model of racial position. *Journal of Personality and Social Psychology*, 112, 696–717.
- Apfelbaum, E. P., Stephens, N. M., & Reagans, R. E. (2016). Beyond one-size-fits-all: Tailoring diversity approaches to the representation of social groups. *Journal of Personality and Social Psychology*, 111, 547–568.
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Gaba, J. A., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143, 1765–1785.
- Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145, 1001–1016.
- Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, 28, 233–248.
- Ross, L., & Nisbett, R. E. (2011). *The person and the situation: Perspectives of social psychology*. London, England: Pinter & Martin.
- Kalev, A., Dobbin, F., & Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, 71, 589–617.
- Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A critical look at evidence from the field and the laboratory. *Annual Review of Psychology*, 60, 339–367.
- Paluck, E. L. (2012). Interventions aimed at the reduction of prejudice and conflict. In L. R. Tropp (Ed.), *Oxford handbook of intergroup conflict* (pp. 179–192). New York, NY: Oxford University Press.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18.
- Weisbuch, M., Pauker, K., & Ambady, N. (2009, December 18). The subtle transmission of race bias via televised nonverbal behavior. *Science*, 326, 1711–1714.
- 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, 7778–7783.
- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, 48, 1267–1278.
- 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, 1227–1274.

32. 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, 452–471.
33. Neel, R., & Shapiro, J. R. (2012). Is racial bias malleable? Whites' lay theories of racial bias predict divergent strategies for interracial interactions. *Journal of Personality and Social Psychology*, 103, 101–120.
34. re:Work Editors (2017, August 7). Use nudges to make your organization better [Blog post]. Retrieved from <https://rework.withgoogle.com/blog/use-nudges-to-make-your-organization-better/>
35. Dobbin, F., Kalev, A., & Kelly, E. (2007). Diversity management in corporate America. *Contexts*, 6(4), 21–27.
36. Wilkins, D. B., & Gulati, G. M. (1996). Why are there so few Black lawyers in corporate law firms? An institutional analysis. *California Law Review*, 84, 493–625.
37. Snyder, K. (2014, August 26). The abrasiveness trap: High-achieving men and women are described differently in reviews. *Fortune*. Retrieved from <https://fortune.com/2014/08/26/performance-review-gender-bias/>
38. McKinsey & Company. (2018). Women in the workplace 2018. Retrieved from https://wiw-assets-production.s3.us-east-1.amazonaws.com/pdfs/women_in_the_workplace_2018_print.pdf
39. Ragins, B. R., Singh, R., & Cornwell, J. M. (2007). Making the invisible visible: Fear and disclosure of sexual orientation at work. *Journal of Applied Psychology*, 92, 1103–1118.
40. Riordan, C. M. (2014, June 5). Diversity is useless without inclusivity. *Harvard Business Review*. Retrieved from <https://hbr.org/2014/06/diversity-is-useless-without-inclusivity>
41. Lai, C. K. (2019, February). *Implicit bias training is teaching*. Paper presented at the meeting of the Society for Personality and Social Psychology, Portland, OR. Retrieved from <https://osf.io/khrcz/>
42. 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.
43. Richeson, J. A., & Shelton, J. N. (2007). Negotiating interracial interactions: Costs, consequences, and possibilities. *Current Directions in Psychological Science*, 16, 316–320.
44. Schultz, J. R., Gaither, S. E., Urry, H. L., & Maddox, K. B. (2015). Reframing anxiety to encourage interracial interactions. *Translational Issues in Psychological Science*, 1, 392–400.
45. Trawalter, S., & Richeson, J. A. (2006). Regulatory focus and executive function after interracial interactions. *Journal of Experimental Social Psychology*, 42, 406–412.
46. Trawalter, S., & Richeson, J. A. (2008). Let's talk about race, baby! When Whites' and Blacks' interracial contact experiences diverge. *Journal of Experimental Social Psychology*, 44, 1214–1217.
47. Brannon, T. N., Carter, E. R., Murdock-Perriera, L. A., & Higginbotham, G. D. (2018). From backlash to inclusion for all: Instituting diversity efforts to maximize benefits across group lines. *Social Issues and Policy Review*, 12, 57–90.
48. DiAngelo, R. (2011). White fragility. *The International Journal of Critical Pedagogy*, 3(3), 54–70.
49. Carter, E. R., & Murphy, M. C. (2015). Group-based differences in perceptions of racism: What counts, to whom, and why? *Social and Personality Psychology Compass*, 9, 269–280.
50. Norton, M. I., & Sommers, S. R. (2011). Whites see racism as a zero-sum game that they are now losing. *Perspectives on Psychological Science*, 6, 215–218.
51. Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, 129, 414–446.
52. Howell, J. L., Collisson, B., Crysel, L., Garrido, C. O., Newell, S. M., Cottrell, C. A., . . . Shepperd, J. (2013). Managing the threat of impending implicit attitude feedback. *Social Psychological and Personality Science*, 4, 714–720.
53. Galinsky, A. D., Todd, A. R., Homan, A. C., Phillips, K. W., Apfelbaum, E. P., Sasaki, S. J., . . . Maddux, W. W. (2015). Maximizing the gains and minimizing the pains of diversity: A policy perspective. *Perspectives on Psychological Science*, 10, 742–748.
54. 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, 1472–1477.
55. Brown, T. N., Tanner-Smith, E. E., Lesane-Brown, C. L., & Ezell, M. E. (2007). Child, parent, and situational correlates of familial ethnic/race socialization. *Journal of Marriage and Family*, 69, 14–25.
56. Hughes, D., Rodriguez, J., Smith, E. P., Johnson, D. J., Stevenson, H. C., & Spicer, P. (2006). Parents' ethnic-racial socialization practices: A review of research and directions for future study. *Developmental Psychology*, 42, 747–770.
57. Jordan, C. H., Spencer, S. J., & Zanna, M. P. (2003). I love me . . . I love me not: Implicit self-esteem, explicit self-esteem, and defensiveness. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *Motivated social perception: The Ontario symposium* (Vol. 9, pp. 117–145). Mahwah, NJ: Erlbaum.
58. Monteith, M. J., Lybarger, J. E., & Woodcock, A. (2009). Schooling the cognitive monster: The role of motivation in the regulation and control of prejudice. *Social and Personality Psychology Compass*, 3, 211–226.
59. Son Hing, L. S., Bobocel, D. R., & Zanna, M. P. (2002). Meritocracy and opposition to affirmative action: Making concessions in the face of discrimination. *Journal of Personality and Social Psychology*, 83, 493–509.
60. 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.
61. Yan, V. X., Bjork, E. L., & Bjork, R. A. (2016). On the difficulty of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General*, 145, 918–935.
62. Albarracín, D., Durantini, M. R., Earl, A., Gunnoe, J. B., & Leeper, J. (2008). Beyond the most willing audiences: A meta-intervention to increase exposure to HIV-prevention programs by vulnerable populations. *Health Psychology*, 27, 638–644.
63. Earl, A., & Nisson, C. (2015). Applications of selective exposure and attention to information for understanding health and health disparities. In R. Scott & S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable,*

- and linkable resource. <https://doi.org/10.1002/9781118900772.etrds0013>
64. Kaiser, C. R., Major, B., Jurcevic, I., Dover, T. L., Brady, L. M., & Shapiro, J. R. (2013). Presumed fair: Ironic effects of organizational diversity structures. *Journal of Personality and Social Psychology*, 104, 504–519.
 65. Ikizer, E. G., & Blanton, H. (2016). Media coverage of “wise” interventions can reduce concern for the disadvantaged. *Journal of Experimental Psychology: Applied*, 22, 135–147.
 66. Duguid, M. M., & Thomas-Hunt, M. C. (2015). Condoning stereotyping? How awareness of stereotyping prevalence impacts expression of stereotypes. *Journal of Applied Psychology*, 100, 343–359.
 67. Emerson, J. (2017, July 17). Don't give up on unconscious bias training—Make it better. *Harvard Business Review*. Retrieved from <https://hbr.org/2017/04/dont-give-up-on-unconscious-bias-training-make-it-better>
 68. Onyeador, I. N., Henderson, K. A., & Shapiro, J. R. (2017, October). Presumed unintentional: Ironic effects of implicit bias framing on perceptions of discrimination. In I. Jurcevic & E. M. Caruso (Chairs), “... Unless the mind is guilty”: How perspective and framing distort judgments of criminal liability. Symposium conducted at the annual conference of the Society of Experimental Social Psychology, Boston, MA.
 69. Daumeyer, N. M., Onyeador, I. N., Brown, X., & Richeson, J. R. (2019). Consequences of attributing discrimination to implicit vs. explicit bias. *Journal of Experimental Social Psychology*, 84, 103812.
 70. Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communications Monographs*, 59, 329–349.
 71. Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. P. Forgas (Ed.), *Studies in emotion and social interaction, second series. Feeling and thinking: The role of affect in social cognition* (pp. 59–82). New York, NY: Cambridge University Press.
 72. Fisher, W. A., Fisher, J. D., & Harman, J. (2003). The information–motivation–behavioral skills model: A general social psychological approach to understanding and promoting health behavior. In J. Suls & K. A. Wallston (Eds.), *Social psychological foundations of health and illness* (pp. 82–106). Hoboken, NJ: Blackwell.
 73. Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54, 493–503.
 74. Wilson, K., Senay, I., Durantini, M., Sánchez, F., Hennessy, M., Spring, B., & Albarracín, D. (2015). When it comes to lifestyle recommendations, more is sometimes less: A meta-analysis of theoretical assumptions underlying the effectiveness of interventions promoting multiple behavior domain change. *Psychological Bulletin*, 141, 474–509.
 75. Jefferson, H., & Lewis, N., Jr. (2018, April 23). Starbucks won't have any idea whether its diversity training works. *Washington Post*. Retrieved from <https://www.washingtonpost.com/news/posteverything/wp/2018/04/23/starbucks-wont-have-any-idea-whether-its-diversity-training-works/>
 76. State v. Gregory, 427 P.3d 621 (Wash. 2018).
 77. Capers, Q., Clinchot, D., McDougale, L., & Greenwald, A. G. (2017). Implicit racial bias in medical school admissions. *Academic Medicine*, 92, 365–369.
 78. Paluck, E. L. (2006). Diversity training and intergroup contact: A call to action research. *Journal of Social Issues*, 62, 577–595.
 79. Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878–902.
 80. Woodzicka, J. A., Mallett, R. K., Hendricks, S., & Pruitt, A. V. (2015). It's just a (sexist) joke: Comparing reactions to sexist versus racist communications. *Humor*, 28, 289–309.
 81. Parker, L. R., Monteith, M. J., Moss-Racusin, C. A., & Van Camp, A. R. (2018). Promoting concern about gender bias with evidence-based confrontation. *Journal of Experimental Social Psychology*, 74, 8–23.
 82. Lewis, N. A., Jr., & Yates, J. F. (2019). Preparing disadvantaged students for success in college: Lessons learned from the preparation initiative. *Perspectives on Psychological Science*, 14, 54–59.
 83. Oyserman, D., & Lewis, N. A. (2017). Seeing the destination and the path: Using identity-based motivation to understand and reduce racial disparities in academic achievement. *Social Issues and Policy Review*, 11, 159–194.
 84. Carter, E. R., & Murphy, M. C. (2017). Consensus and consistency: Exposure to multiple discrimination claims shapes Whites' intergroup attitudes. *Journal of Experimental Social Psychology*, 73, 24–33.
 85. Dovidio, J. F., Gaertner, S. L., Dittmann, R. K., & West, T. V. (2012). Intergroup relations in post-conflict contexts: How the past influences the present (and future). In K. J. Jonas & T. A. Morton (Eds.), *Restoring civil societies: The psychology of intervention and engagement following crisis* (pp. 135–155). Malden, MA: Wiley-Blackwell.
 86. Eibach, R. P., & Ehrlinger, J. (2006). “Keep your eyes on the prize”: Reference points and racial differences in assessing progress toward equality. *Personality and Social Psychology Bulletin*, 32, 66–77.
 87. Nelson, J. C., Adams, G., & Salter, P. S. (2013). The Marley hypothesis: Denial of racism reflects ignorance of history. *Psychological Science*, 24, 213–218.