

# A systematic review of diversity, equity, and inclusion and antiracism training studies: Findings and future directions

Monica L. Wang,<sup>1,2,\*</sup> Alexis Gomes,<sup>3</sup> Marielis Rosa,<sup>3,4</sup> Phillipe Copeland,<sup>5</sup> Victor Jose Santana<sup>6</sup>

<sup>1</sup>Department of Community Health Sciences, Boston University School of Public Health, Boston, MA, USA;

<sup>2</sup>Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA, USA;

<sup>3</sup>Boston University Center for Antiracist Research, Boston, MA, USA;

<sup>4</sup>Boston University College of Arts and Sciences, Boston, MA, USA;

<sup>5</sup>Boston University School of Social Work, Boston, MA, USA

<sup>6</sup>VJS Consulting, Boston, MA, USA

\*Correspondence address. Chair of the Narrative Office, Boston University Center for Antiracist Research, Associate Professor, Boston University School of Public Health, 801 Massachusetts Ave, Boston, MA 02118, USA. E-mail: [mlwang@bu.edu](mailto:mlwang@bu.edu)

## Abstract

A growing number of organizations are prioritizing diversity, equity, and inclusion (DEI) and antiracism in the workplace, including investing resources in DEI or antiracism training. However, such trainings vary widely in curriculum, objectives, delivery, and evaluation, with little known about the efficacy of existing trainings. The aim of this systematic review is to evaluate training characteristics, measures, and results of peer-reviewed studies (published between 2000 and 2022) testing DEI or antiracism trainings. Studies were identified using Google Scholar, JSTOR, and a university library database. Key search terms included “diversity, equity, and inclusion training”; “antiracism training”; and “effect,” “impact,” “outcome,” or “evaluation.” The search yielded  $N = 15$  DEI training studies and  $N = 8$  antiracism training studies. The majority of studies (75% of antiracism training; 66.6% of DEI training) utilized a one-time training session. Content, objectives, measures, and impact varied widely across studies. Randomized designs were uncommon (13%), and over 70% of studies had majority female participants. Findings highlight several strategies to advance the field of DEI and antiracism training, such as shifting curriculum from targeting individual knowledge to supporting behavioral and organizational change, providing longitudinal training, standardizing outcomes of interest, and implementing rigorous evaluation methods.

## Lay summary

Despite the increased number of organizations dedicating resources to diversity, equity, and inclusion training and/or antiracism training, little is known about which strategies yield successful results. This systematic review synthesizes research findings on diversity, equity, and inclusion and antiracism training studies over the past two decades and presents research and practice-based recommendations for how to move the field forward.

**Keywords:** diversity, equity, and inclusion (DEI); antiracism; training; efficacy

## Implications

**Practice:** Diversity, equity, and inclusion (DEI) and antiracism trainings may benefit from including longitudinal training design, targeting skill building and organizational development as core parts of the curriculum, and prioritizing behavioral and organizational change as outcomes of interest.

**Policy:** Investments in research on DEI and antiracism trainings are needed to identify best practices for cultivating equitable and inclusive work environments.

**Research:** Future research is needed that integrates community-engaged approaches in the design, implementation, evaluation, and dissemination of DEI and antiracism trainings. Additionally, the use of more rigorous study design methods and longer follow-up assessments can help identify efficacious training strategies.

## Introduction

Advancing diversity, equity, and inclusion (DEI) and anti-racism in the workforce has increased in priority across industries and sectors, including health care and higher education, over the past few decades and in particular over recent years. Following the racial reckoning of 2020 in the USA and globally, the number of DEI-related job positions increased by 60% in the USA [1] and 86% of large financial companies surveyed stated intention to increase investments in DEI training [2]. In 2021, the USA spent nearly \$3.8 billion dollars in DEI-related training, accounting for 45.5% of the total global spending on DEI [3]. By 2026, global DEI-related spending is expected to reach \$15.4 billion [3].

Despite the growing number of organizations expressing commitment and dedicating resources to DEI and/or anti-racism training, little is known about the efficacy of such trainings (T2 of the translational research spectrum) [4]. Increasing our understanding of what works and what does not will allow us to better identify best practices, avoid investing in inefficacious strategies, and innovate new, community-engaged research approaches that are better positioned to make an impact [5]. The need for novel and impactful DEI and anti-racism trainings is great, as inequities in the workplace continue to persist. Across industries, pay disparity continues to affect employees based on race [6], gender [7], age [8], disability [9], and sexual orientation [10]. A recent Gallup poll results revealed that 61% of Latine employees and 75% of Black employees in the USA experienced ethnicity or race-based discrimination in the last year [11]. In 2021, over 47 million employees voluntarily quit their jobs, and subsequent analysis revealed that toxic culture was the primary driver of attrition, with the failure of organizations to promote DEI among the top characteristics of an unhealthy work culture [12].

Studies to date evaluating such trainings have yielded inconsistent findings, with some demonstrating change in knowledge and attitudes associated with training participation [13, 14] while others show null results [15, 16]. The mixed findings may partially be attributed to the wide heterogeneity of training characteristics (e.g. goals, curriculum, duration), study participants and settings, and outcomes measured. Given the wide range of training and study characteristics, research to date on the efficacy of DEI and anti-racism trainings has been unable to reach a consensus regarding how such trainings ought to be designed, measured, implemented, and evaluated to advance equity in the workplace.

While the definition and purpose of DEI and anti-racism trainings has varied over time and context, we define DEI training as the intentional provision and application of education (in various forms) to “recognize and address how power, privilege, and how society affects our personal identities,” including race, socioeconomic status, sexual orientation, gender, etc [17]. We defined anti-racism training as the intentional provision and application of education to help individuals and organizations “understand systemic racism, along with the practices or structures that perpetuate it” [17]. The goal of this systematic review was to evaluate the training content and strategies, theoretical frameworks, delivery modality and duration, participant characteristics, measures, and results of peer-reviewed research studies (published between 2000 and 2022) of DEI and anti-racism trainings tested in work settings. As DEI and anti-racism trainings are needed and applicable

across all industries and sectors, including medicine, public health, and higher education, we did not limit our search to studies conducted in health care or health-related settings.

## Methods

### Data sources, study selection, and data abstraction

The systematic review was conducted using three online databases: Google Scholar, JSTOR, and a university libraries database, which included relevant databases such as PsycINFO, Wiley Online Library, SpringerLink, ScienceDirect, and others. Key search terms for DEI training studies included “diversity, equity, and inclusion training” and [“effect,” “impact,” “outcome,” or “evaluation”]. Key search terms for anti-racism training studies included “anti-racism training” and [“effect,” “impact,” “outcome,” or “evaluation”].

Inclusion criteria for the studies of interest were: (i) original, empirical study that tested and measured the effect(s) of a DEI or an anti-racism training; (ii) peer-reviewed publication between 1 January 2000 and 31 December 2022; and (iii) published in the English language. One coauthor conducted a search for DEI training studies and a separate search for anti-racism training studies based on the search terms and engines specified above. Article titles and abstracts were screened based on inclusion criteria. Articles that met inclusion criteria based on the abstract review were selected for a full-text review. Articles that met all criteria were included for review.

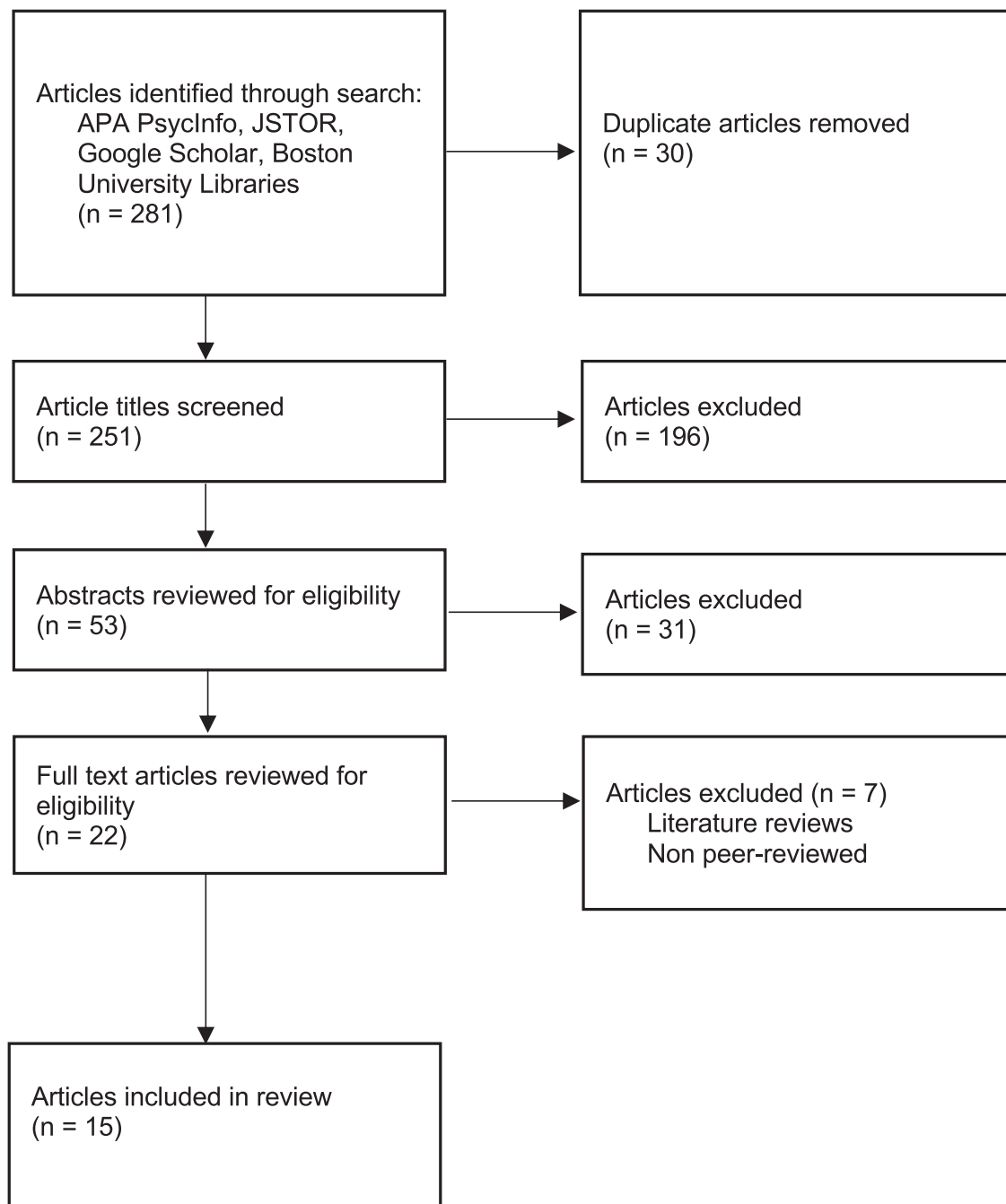
Data abstracted included sample characteristics (race, gender, and other reported sociodemographics of participants, number of participants), training characteristics [theoretical frameworks that informed the training (if any), training format, delivery, topics/curriculum, duration, and setting], study design (e.g. quasi-experimental vs. randomized, use of a comparison group, follow-up duration), outcomes assessed (e.g. knowledge, attitudes, skills, behavior change) and survey instruments used, and training effect(s) on measured outcomes of interest. One reviewer first abstracted data of interest for each article and an additional reviewer conducted independent data validations. The training topics were gleaned from the descriptions of the interventions in the studies’ methods section; two independent reviewers reviewed the methods section of all the studies and described the training topics. Discrepancies were reviewed and resolved by consensus with a third reviewer.

## Results

As DEI and anti-racism training are conceptually distinct, we present the results of DEI and anti-racism training studies reviewed separately and with patterns observed across both types of studies. Studies included for the systematic review based on the DEI training search methods did not overlap with studies identified via the anti-racism training search methods.

### DEI training search results

The literature search for DEI training studies yielded a total of 251 nonduplicate articles (Figure 1). Of those, 198 articles were excluded based on the title review.  $N = 31$  studies were excluded during the abstract review process, and  $N = 7$  studies were excluded during the full-text review, resulting in a total of 15 articles eligible for inclusion. Table 1 presents



**Figure 1** Flow diagram of study article inclusion and exclusion: DEI training studies.

a summary of characteristics across DEI training studies included in this review.

#### DEI training studies: Participant, setting, and training characteristics

The number of participants in DEI training studies in our review varied widely from 30 to 3016 with 7 studies (46.7%) having sample sizes of 100 participants or less. Most studies (66.7%) had participants that were majority (60% or more) female, and 46.7% of studies had participant samples that were majority (60% or more) White. The majority of DEI training studies took place in higher education (73.4%), followed by corporate (13.3%) and health care (13.3%) settings.

The majority of trainings (80%) were conducted in-person by trained facilitators, with two studies' training activities occurring online. No studies utilized a hybrid delivery approach.

Studies varied widely in the duration, frequency, and curriculum of DEI training. More than half of the studies (53.3%) used only one-time training sessions that ranged from 1-hour to 3-day sessions. The remaining studies (46.7%) implemented multiple training sessions spaced out over time, ranging from 6 weeks to 10 months. Two-thirds of DEI training studies (66.7%) identified one or more theories to inform their training, including social identity theory [33], self-determination theory [34], and diversity management theory [35]. See Supplementary Appendix 1 for a complete list and summary

**Table 1** Characteristics and results of DEI training studies published between 2000 and 2022 ( $N = 15$ )

| Study article                         | Sample and setting characteristics   | Training characteristics  | Measures  | Analysis and results  |
|---------------------------------------|--|---|---|---|
| 1. Castillo <i>et al.</i> , 2011 [18] | $N = 84$ participants<br>Demographics:<br>- Gender: 80% female<br>- Race: 77.4% White, 3.6% Black, 6% Asian, 6% Latino, 7% biracial<br>Setting: higher education | Theory: adverse racism theory, feminist theory<br>Modality: in-person<br>Duration: weekly 3-hour course for 15 weeks<br>Facilitator characteristics: all were professors; 2 were Latina, 3 were White (1 female, 2 male)<br>Topics covered: history, cultural customs, traditions, and values, feminism | Study design: course evaluations analyzed from 10 courses (5 multicultural, 5 counseling foundations)<br>Outcomes measured: knowledge, attitudes, skills, self-awareness<br>Instruments used: Multicultural Counseling Inventory, Race Implicit Association Test, study survey    | Assessment time points: Pre and post-training<br>Analysis: multilevel modeling<br>Effect(s): null improvements in knowledge and skills, statistically significant improvement in awareness  |
| 2. Cavaleros <i>et al.</i> [19]       | $N = 382$ participants<br>Demographics:<br>- Gender: 174 women, 208 men<br>- Race: 78 non-white, 304 white<br>Setting: corporate                                 | Theory: cognitive theory<br>Modality: in-person<br>Duration: One 2-day session<br>Facilitator characteristics: NR<br>Topics covered: racism, sexism, classism, prejudice, discrimination, and bias  | Study design: Solomon four-group design (see page 55 of Cavaleros <i>et al.</i> )<br>Outcomes measured: awareness, job satisfaction<br>Instruments used: study survey, Minnesota Job Satisfaction Questionnaire   | Assessment time points: pre- and 1 month post-training<br>Analysis: $2 \times 2$ ANOVA<br>Effect(s): null improvements in both awareness of diversity and job satisfaction  |
| 3. Celik <i>et al.</i> [20]           | $N = 30$ participants<br>Demographics:<br>- Gender: 60% female<br>- Age range: 25–60<br>Setting: health care   | Theory: Deming cycle<br>Modality: in-person<br>Duration: Four 4-hour sessions<br>Facilitator characteristics: NR<br>Topics covered: levels of diversity (individual, symbolic, institutional), advantages of diversity for professionals, applying diversity in the institution                         | Study design: Quantitative one-group pretest–posttest and qualitative post interviews<br>Outcomes measured: awareness, knowledge, behavior<br>Instruments used: study survey, interviews  | Assessment time points: pre- and post-training; 3 years after (interview)<br>Analysis: statistical comparisons measured using Wilcoxon signed rank test<br>Effect(s): significant improvements in awareness; null effects on knowledge, attitudes, and behavior   |
| 4. Chang <i>et al.</i> [21]           | $N = 3016$ participants<br>Demographics:<br>- Gender: 61.5% male<br>Setting: corporate   | Theory: NR<br>Modality: online<br>Duration: One 1-hour session<br>Facilitator characteristics: NR<br>Topics covered:<br>- Gender-bias training: gender-bias, stereotyping<br>- General-bias training: gender, race, and sexual orientation bias, stereotyping   | Study design: Random assignment to 1 of 3 experimental conditions: gender-bias training, general-bias training, or a control training<br>Outcomes measured: attitudes and behavior (workplace decisions)<br>Instruments used: Implicit Association Test, study survey (validated) | Assessment time points: pre- and post-training, 3 weeks, 6 weeks, 14 weeks (follow-ups for gender-bias group only)<br>Analysis: controlled ordinary least-squares regressions with interactions terms<br>Effect(s): significant effects on attitudes toward women, acknowledgment of gender bias, and intentions to engage in inclusive workplace behaviors toward women (non-US participants only). Null effects at follow-up. Significant effects in acknowledgement of racial bias and number of racial minorities recognized for excellence (US participants) |

Table 1. Continued

| Study article                          | Sample and setting characteristics  | Training characteristics  | Measures  | Analysis and results   |
|--|---|---|---|--|
| 5. Davis <i>et al.</i> [22]            | N = 615 participants<br>Demographics: NR. All participants were first-year medical students at the University of California, San Francisco Medical School<br>Setting: higher education  | Theory: NR<br>Modality: in-person<br>Duration: One 2-day session<br>Facilitator characteristics: received 1-day training to prepare<br>Topics covered: health disparities, communication skills, privilege, power, socio-cultural history, and microaggressions   | Study design: Mandatory one-group posttest<br>Outcomes measured: knowledge, skills, attitudes<br>Instruments used: study survey   | Assessment time points: 4 weeks post-training<br>Analysis: qualitative data analyzed by 2 authors<br>Effects: Null results in improvement in knowledge   |
| 6. Ehrke <i>et al.</i> [23]            | N = 62 participants<br>Demographics:<br>- Gender: 84% female [Experiment 1]; 77% female [Experiment 2]<br>- Age: ages 19–31<br>Setting: higher education  | Theory: ingroup projection model<br>Modality: in-person<br>Duration: 2 hours [Experiment 1]; 1 day [Experiment 2]<br>Facilitator characteristics: NR [Experiment 1]; 2 experienced trainers (1 male, 1 female), prepared with training manual and “train the trainer” session [Experiment 2]<br>Topics covered: gender stereotypes [Experiment 1], stereotypes, social categorization, discrimination, general diversity [Experiment 2]<br>Theory: training theory and cross-cultural training models | Study design:<br>2 (condition: diversity or control) × 3 (time: 1, 2, 3) design [Experiment 1]; cross-over design with three measurement timepoints [Experiment 2]<br>Outcomes measured: perceived diversity, self-typicality attitudes<br>Instruments used: study survey | Assessment time points:<br>pre- and post-training, 1 month [Experiment 1]<br>Analysis: ANOVAs<br>Effect(s): significant effect in perceived diversity of students in diversity-intervention condition, self-typicality, and positive attitudes toward gender outgroup; significant increase in perceived diversity and reduced ambivalent sexism, age bias, sexist attitudes. Null improvements in anti-gay attitudes. |
| 7. Fischer [24]                        | N = 49 participants<br>Demographics:<br>- Average age: 22<br>- Gender: 78.4% female<br>- Race: 44.7% non-white<br>Setting: higher education   | Theory: training theory and cross-cultural training models<br>Modality: in-person<br>Duration: weekly 2-hour lectures for 6 weeks<br>Facilitator characteristics: trained facilitators<br>Topics covered: cultural diversity, culture theory, dimensions of culture, cultural competence, intercultural training  | Study design: Two-group pretest–posttest<br>Outcomes measured: cultural intelligence, open-mindedness, cultural essentialism<br>Instruments used: Multicultural personality Questionnaire, study survey   | Assessment time points:<br>pre- and post-training<br>Analysis: within-subject ANOVA and regression analysis<br>Effect(s): significant improvement in cultural essentialism, null effect in open-mindedness   |
| 8. Harrison-Bernard <i>et al.</i> [25] | N = 55 participants<br>Demographics:<br>- Gender: 67% female, 31% male, 2% no response<br>- Race: 12% Black, 14% Asian, 7% Hispanic/Latinx, 5% no response, 62% White<br>- Age: 5% 25–34, 29% 35–44, 38% 45–54, 19% 55–64, 7% 65+, 2% NR<br>Setting: higher education | Theory: NR<br>Modality: in-person<br>Duration: One 3-hour workshop<br>Facilitator characteristics: facilitators were required to use foundational materials to inform the training (see Table 1 on pg. 287 of Harrison-Bernard <i>et al.</i> )<br>Topics covered: importance of a diverse workforce, inclusion, stereotypes, implicit bias, cultural competency   | Study design: one-group pretest–posttest<br>Outcomes measured: awareness, knowledge, self-perceptions<br>Instruments used: study survey   | Assessment time points: pre- and post-training<br>Analysis: paired <i>t</i> test<br>Effects: Statistically significant increase in self-perceptions regarding the importance of improving diversity, recognizing bias and stereotypes, and cultural awareness. Statistically significant improvement in knowledge of terms (color-blind racial attitudes, tokenism, failure to differentiate, status leveling).        |

**Table 1.** Continued

| Study article                  | Sample and setting characteristics   | Training characteristics  | Measures  | Analysis and results   |
|--------------------------------|--|---|---|--|
| 9. Holladay and Quinones [26]  | <p>N = 191 participants</p> <p>Demographics:</p> <ul style="list-style-type: none"> <li>- Gender: 84 men, 107 women</li> <li>- Race: 108 white, 8 Black, 21 Hispanic, 37 Asian, 17 other</li> </ul> <p>Setting: higher education</p>   | <p>Theory: NR</p> <p>Modality: online</p> <p>Duration: One 1-hour session</p> <p>Facilitator characteristics: facilitators ranged in sex and race (White men, White women, Black men, Black women)</p> <p>Topics covered: benefits of diversity</p>   | <p>Study design: quasi-experimental with 6 conditions; <math>2 \times 2 \times 2</math> focus (training focus: similarities vs. differences) <math>\times</math> (trainer sex) <math>\times</math> (trainer race: White vs. Black)</p> <p>Outcomes measured: reactions, knowledge, attitudes, self-efficacy</p> <p>Instruments used: study survey</p> | <p>Assessment time points: pre- and post-training, no data reported for follow-ups</p> <p>Analysis: mediation analysis</p> <p>Effect(s): Trainees who participated in a diversity training program focused on similarities expressed significantly lower levels of expected backlash and performed better on the situational judgment test of conflict resolution skills, similarity-focused program significantly mitigated the effects of trainer race and sex on trainee reactions.</p> |
| 10. Hostager and De Meuse [27] | <p>N = 302 participants</p> <p>Demographics:</p> <ul style="list-style-type: none"> <li>- Gender: 47% female, 53% male</li> <li>- Race: 94% white, 3% Asian or Pacific Islander, 2% Hispanic, 1% American Indian or Alaskan Native</li> </ul> <p>Setting: higher education</p> | <p>Theory: NR</p> <p>Modality: in-person</p> <p>Duration: one semester course</p> <p>Facilitator characteristics: NR</p> <p>Topics covered:</p> <ul style="list-style-type: none"> <li>- Course 1 (Diversity in the Workplace): workforce demographics and diversity, prejudice, discrimination, racism, legislation, gender diversity and harassment, diversity trends</li> <li>- Course 2 (Organizational Change and Development): strategic forms of organizational change, organizational change theory, employee empowerment, personality types</li> <li>- Course 3 (Strategic Management in Global Business Environment): maximizing company performance through knowledge of organizational structure, cultural, politics, and change, theories of ethics, stakeholder management, market trends.</li> </ul> | <p>Study design: Quasi-experiment with three conditions and pretest and posttest. Participants enrolled in one of three courses.</p> <p>Outcomes measured: reactions (emotional and behavioral), judgements, personal consequences, organizational outcomes</p> <p>Instruments used: study survey, Reaction-to-Diversity Inventory</p>                | <p>Assessment time points: pre- and post-training</p> <p>Analysis: ANOVA, MANCOVA</p> <p>Effect(s): significant effect of courses describing both positive and negative aspects of diversity</p>   |

Table 1. Continued

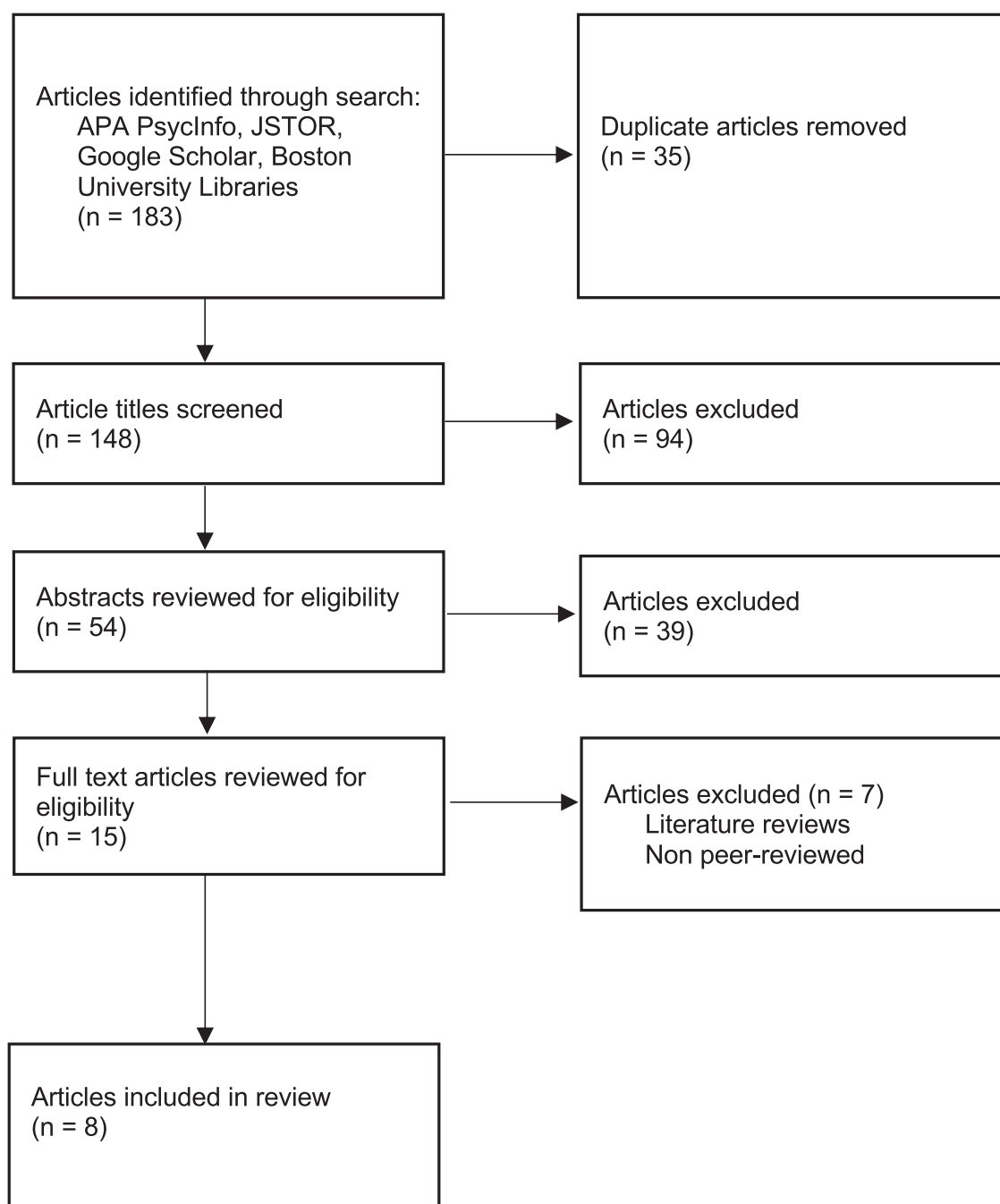
| Study article                   | Sample and setting characteristics  | Training characteristics  | Measures  | Analysis and results   |
|---------------------------------|---|---|---|--|
| 11. Lindsey <i>et al.</i> [28]  | <p>N = 118 participants</p> <p>Demographics:</p> <ul style="list-style-type: none"><li>- Gender: 61 % female, 39 % male</li><li>- Race: 62 % White, 12 % Asian, 10 % Black, 6 % Hispanic, 10 % other</li></ul> <p>All incoming freshman at a southern university in 2006.</p> <p>Setting: higher education</p>  | <p>Theory: diversity management theory and social identity theory</p> <p>Modality: in-person</p> <p>Duration: One 2-hour training</p> <p>Facilitator characteristics: NR</p> <p>Topics covered: varied dependent on group;</p> <ul style="list-style-type: none"><li>- Goal setting: setting goals related to diversity training</li><li>- Stereotype: discrediting stereotypes</li><li>- Perspective taking: challenges faced by marginalized groups</li></ul>   | <p>Study design: 3x2 design. Participants randomly assigned to 1 of 3 diversity training methods (perspective taking vs. goal setting vs. stereotype discrediting) and 1 of 2 groups (African-Americans vs. LGBT individuals).</p> <p>Outcomes measured: attitudes, motivation, empathy</p> <p>Instruments used: 7-item Modern Racism Scale; 10-item Attitudes Toward Lesbians and Gay Male Scale; 10-item scale from International Personality Item Pool; study survey</p> | <p>Assessment time points: Pre and post-training; 0, 3, 6 months</p> <p>Analysis: regression analyses and SPSS macros</p> <p>Effect(s): Null results in the stereotype discrediting and goal setting conditions for both African-American and LGBT groups; statistically significant results in the perspective taking condition for the LGBT group.</p> |
| 12. O'Connor <i>et al.</i> [29] | <p>N = 44 participants</p> <p>Demographics: Not reported (NR); all participants were faculty and staff at University of Washington School of Nursing</p> <p>Setting: higher education</p>   | <p>Theory: NR</p> <p>Modality: in-person</p> <p>Duration: One 3-day training</p> <p>Facilitator characteristics: professional social justice workers (no specific training given) and DEI faculty mentors who received a 5-hour training on inclusive teaching practices</p> <p>Topics covered: awareness building, creating inclusive environments for diverse learners, managing difficult conversations on racism, incorporating DEI into curriculum, normalizing and handling mistakes with humility</p>  | <p>Study design: one-group pretest–posttest</p> <p>Outcomes measured: behavior change</p> <p>Instruments used: study survey, University of Washington course evaluation form</p>  | <p>Assessment time points: pre- and post-training</p> <p>Analysis: nonparametric Kolmogorov-Smirnov test for equality of distributions, paired <i>t</i> tests</p> <p>Effects: statistically significant<sup>a</sup> increases in DEI-related teaching self-efficacy</p>  |
| 13. Powell <i>et al.</i> [30]   | <p>N = 74 participants</p> <p>Demographics:</p> <ul style="list-style-type: none"><li>- Gender: 65 % female, 34 % male, 1 % prefer to self-describe (queer)</li><li>- Race/ethnicity: 1 % American Indian/Alaskan Native, 31 % Asian/Asian American, 22 % Black, 20 % Hispanic/Latino/Spanish origin, 22 % non-Hispanic White, 5 % other</li></ul> <p>Setting: higher education</p> | <p>Theory: self-determination theory</p> <p>Modality: in-person</p> <p>Duration: monthly 2-hour sessions over the course of 10 months</p> <p>Facilitator characteristics: no training received, facilitators were selected from an applicant pool</p> <p>Topics covered: importance of diversity, implicit bias, microaggressions, emotional intelligence, growth and fixed mindsets, privilege and allyship in the workplace, promoting underrepresented trainees in leadership, principles of compassionate leadership, “inclusive diversity” culture, structural racism in medicine, imposter syndrome, social justice in medicine</p> | <p>Study design: LEAD program (longitudinal, single institution program) over 3 cohorts, from 2017 to 2018, 2018 to 2019, and 2019 to 2020.</p> <p>Outcomes measured: self-efficacy, actual and planned behavior change</p> <p>Instruments used: IRB-approved, de-identified retrospective pre- and post-surveys</p>  | <p>Assessment time points: Pre and post-training; annual follow-ups (no data reported)</p> <p>Analysis: descriptive statistics, 2-tailed paired <i>t</i> tests and qualitative analysis. Hand-coded using content analysis</p> <p>Effect(s): statistically significant improvement in self-efficacy for all LEAD sessions’ learning objectives</p>       |

**Table 1.** Continued

| Study article                  | Sample and setting characteristics   | Training characteristics   | Measures  | Analysis and results  |
|--------------------------------|--|--|---|---|
| 14. Rawski and Conroy [31]     | N = 194 (N = 53 in Experiment 1 and N = 141 in Experiment 2)<br>Experiment 1 demographics:<br>- Gender: 74% female<br>- Race: 92% White<br>- Average age: 41.5<br>Experiment 2 demographics:<br>- Gender: 75% female<br>- Race: 81% White<br>- Average age: 43.3<br>Setting: higher education<br>N = 21 participants | Theory: social identity theory<br>Modality: in-person<br>Duration:<br>- Experiment 1: voluntary attendance of 14, one 4-hour diversity training workshops<br>- Experiment 2: one 1-hour session<br>Facilitator characteristics: described as subject matter experts<br>Topics covered:<br>Experiment 1: cultural competency, microaggressions, unconscious bias, power, privilege, and oppression; Experiment 2: sexual harassment<br>Theory: NR<br>Modality: NR<br>Duration: monthly 1-hour sessions from August 2020 to June 2022<br>Facilitator characteristics: no training received, session speakers were sought out on the basis of lived experience<br>Topics covered: Race and ethnicity (identity, implicit and explicit bias, race and economics, racial representation in health professions, micro and macro aggressions, racial equity), gender/gender identity (LGBTQI+ definitions, safe zone project training, health disparities for trans individuals, gender differences in healthcare experiences and health professional training), religion (myths and truths around religion-driven medical decision making); social determinants of health. | Study design: Quasi-experiment with two conditions. Experiment 1 was a multi-session, general diversity training; Experiment 2 was a mandatory, one-session sexual harassment training. Outcomes measured: knowledge, motivation<br>Instruments used: study survey, Organizational Identification Scale | Assessment time points: pre- and post-training<br>Analysis: logistic regression analysis<br>Effect(s): statistically significant higher effect of learning in mandatory training [Experiment 2] but better reactions in voluntary training [Experiment 1] |
| 15. Tillman <i>et al.</i> [32] | Demographics:<br>- Gender: 23.8% male, 71.4% female, 4.8% nonbinary<br>- Race: 14.3% Asian, 19.1% Black, 4.8% Pacific Islander, 57.1% White, 4.8% multi-racial<br>- Age: 4.8% 18–24, 90.48% 25–29, 0% 30–34, 4.8% 35+<br>Setting: health care  |  | Study design: longitudinal DEI curriculum<br>Outcomes measured: knowledge, awareness, satisfaction<br>Instruments used: study survey  | Assessment time points: post-training only<br>Analysis: NR<br>Effect(s): null improvements in knowledge and awareness   |

NR, not reported.

<sup>a</sup> Statistically significant refers to  $P < .05$ .



**Figure 2** Flow diagram of antiracism training study article inclusion and exclusion.

of theories. Training topics varied, with the most common topic being bias (discussed in 42.9% of DEI training studies), stereotypes (26.7%), and cultural competency (20%), among others. Nine of the 15 DEI studies (60%) included training on topics other than race and racism, most commonly sexism, classism, homophobia and transphobia.

### DEI training study design and measures

Very few studies (20%) utilized randomized conditions, and even fewer (6.7%) utilized a control group. All studies administered a pre and immediate post-training survey to measure outcomes. Of the four studies that collected longer-term follow-up data, only two reported results (follow-up ranged from 1 month to 3 years). Measured outcomes varied widely across

studies. The most frequently measured outcome was content knowledge (53.3%), with most studies assessing knowledge, awareness, and attitudes (e.g. motivation, self-efficacy). Seven out of the 15 studies also measured participants' skills or actionable change, such as conflict resolution or decisions. Over half of the studies (53.3%) developed their own study survey for assessing outcomes, with the other half using established inventories including the Race Implicit Association Test [34] and the Organizational Identification Scale [36].

### DEI training impact

Twelve out of the 15 DEI training studies (80%) reported statistically significant results for one or more measured outcomes. In general, several studies reported significant

**Table 2** Characteristics and results of antiracism training studies published between 2000 and 2022 (*N* = 8)

| Study article                 | Sample and setting characteristics   | Training characteristics  | Measures   | Analysis and results   |
|-------------------------------|--|---|--|--|
| 1. Abramovitz and Blitz [37]  | <p><i>N</i> = 875 participants</p> <p>Demographics:</p> <ul style="list-style-type: none"> <li>- Gender: 77.7% female, 22.3% male</li> <li>- Race: 57.8% White, 38.8% person of color, 3.5% other</li> <li>- Age: 31.3% 18–35, 36.8% 36–55, 31.9% 56+</li> </ul> <p>Setting: social work</p> | <p>Theory: participatory action model</p> <p>Modality: in-person</p> <p>Duration: one 2.5 day workshop</p> <p>Facilitator characteristics: facilitated by multi-cultural team</p> <p>Topics covered: history of racism in USA, individual attitudes and knowledge, internalized oppression and privilege, institutional racism</p>  | <p>Study design: one-group pretest and follow-up</p> <p>Outcomes measured: knowledge and attitudes, engagement in job-related racial equity activities, view of organizations' progress toward racial equity</p> <p>Instruments used: study survey</p>                       | <p>Assessment time points: post-training (2–7 years after)</p> <p>Analysis: frequency distributions reported, non-parametric tests used for analysis</p> <p>Effect(s): statistically significant improvements in knowledge, attitudes, and engagement.</p>   |
| 2. Johnson <i>et al.</i> [38] | <p><i>N</i> = 462 participants</p> <p>Demographics:</p> <ul style="list-style-type: none"> <li>- Gender: 78% women, 22% men</li> <li>- Race: 57% White, 36% Black, 6% other</li> <li>- Age: range of 21–83, average 42</li> </ul> <p>Setting: community service providers</p>                | <p>Theory: cultural competency training theory</p> <p>Modality: in-person</p> <p>Duration: 1 training over 2.5 days</p> <p>Facilitator characteristics: 20+ years experience in antiracism education and organizing</p> <p>Topics covered: foundations of race and racism, impact of racism, approaches to community organizing, leadership development, accountability to communities, creation of networks, learning from history, combating racism methods</p> | <p>Study design: one-group, pretest–posttest</p> <p>Outcomes measured: anticipated transfer of knowledge and skills, knowledge of key concepts, attitudes toward race</p> <p>Instruments used: study survey, Color-Blind Racial Attitudes Scale</p>                          | <p>Assessment time points: pre- and post-training</p> <p>Analysis: independent samples <i>t</i>-tests and one-way ANOVAs, paired samples <i>t</i>-tests and repeated measures ANOVAs, bivariate correlations</p> <p>Effect(s): significantly higher satisfaction among participants of color than white participants; significant increase in knowledge; significant decrease in color-blind racial attitudes.</p> |
| 3. Kingett <i>et al.</i> [39] | <p><i>N</i> = 109 participants</p> <p>Demographics:</p> <ul style="list-style-type: none"> <li>- Gender: 74% male, 26% female</li> </ul> <p>Setting: secondary education</p>   | <p>Theory: NR</p> <p>Modality: online</p> <p>Duration: One day (# hours not specified)</p> <p>Facilitator characteristics: training facilitated by Show Racism the Red Card, antiracism educational charity</p> <p>Topics covered: history and effects of racism, stereotypes, racism prevention, prejudice, discrimination</p>   | <p>Study design: one-group pretest–posttest</p> <p>Outcomes measured: knowledge (recognizing racism, perceptions of racial attitudes), awareness, confidence, acceptance/rejection of discrimination, response to racist incidents</p> <p>Instruments used: study survey</p> | <p>Assessment time points: pre- and post-training</p> <p>Analysis: NR</p> <p>Effect(s): statistically significant improvement in knowledge, awareness, confidence, and acceptance of discrimination. Null improvement in response to racist incidents</p>  |
| 4. Loe <i>et al.</i> [40]     | <p><i>N</i> = 77 participants</p> <p>Demographics: NR</p> <p>Setting: health care</p>  | <p>Theory: Kern's 6 steps of curriculum development, self-determination theory</p> <p>Modality: in-person</p> <p>Duration: One 3-hour workshop</p> <p>Facilitator characteristics: NR</p> <p>Topics covered: racial awareness, child development, impact of racism on child development, behavior, and health, addressing racism</p>  | <p>Study design: one-group pretest–posttest</p> <p>Outcomes measured: knowledge, attitudes, skills (comfort and preparation for counseling patients, offering resources, etc.)</p> <p>Instruments used: study survey</p>   | <p>Assessment time points: pre- and post-training</p> <p>Analysis: paired <i>t</i> test</p> <p>Effect(s): significant improvement in all measured outcomes</p>   |

Table 2. Continued

| Study article                 | Sample and setting characteristics  | Training characteristics  | Measures  | Analysis and results  |
|-------------------------------|---|---|---|---|
| 5. Martin and Baxter [41]     | N = 50 participants<br>Demographics:<br>- Race: 100% White<br>- Gender: 24% male, 76% female<br>- Age: range from 30 to 60, average age of 47<br>Setting: secondary education | Theory: 5 stage model of White Racial Identity Development<br>Modality: in-person<br>Duration: 3 hours/week over 14 weeks<br>Facilitator characteristics: NR<br>Topics covered: principles of antiracism, applying antiracism   | Study design: one-group pretest–posttest<br>Outcomes measured: stages of the White racial identity development model (contact, disintegration, reintegration, pseudo-independence, autonomy)<br>Instruments used: White Racial Identity Attitude Survey, study survey | Assessment time points: pre- and post- training<br>Analysis: confirmed reliability with difference approach<br>Effect(s): significant decrease in disintegration (positive result), significant improvement in pseudo-independence and autonomy stages. No changes in contact and reintegration stages. |
| 6. Santoro <i>et al.</i> [42] | N = 42 participants<br>Demographics:<br>- Race: 55% White, 32% Black, 9% Asian, 4% other<br>- Gender: 91% female<br>- Age: average age 30<br>Setting: healthcare (nutrition)  | Theory: cultural humility<br>Modality: in-person<br>Duration: one, 3-hour training<br>Facilitator characteristics: NR, led by Afro-Latinx antiracism educator and White midwife with public health research experience<br>Topics covered: racism, privilege, white supremacy, implicit bias and cultural humility, socio-ecological models of behavior change | Study design: one-group pretest–posttest<br>Outcomes measured: awareness (of systematic racism, bias, privilege), confidence, skills (identifying and addressing racism)<br>Instruments used: study survey  | Assessment time points: pre- and post-training, 6 months post-training<br>Analysis: 2-sided z-tests<br>Effect(s): significant and sustained improvements in awareness and skills, significant but unsustained improvements in confidence  |
| 7. Webb and Sergison [43]     | N = 48 participants<br>Demographics: NR<br>Setting: child health services   | Theory: NR<br>Modality: in-person<br>Duration: NR<br>Facilitator characteristics: NR<br>Topics covered: stereotyping, empathy, communication, attitudes, cultural conditioning, how racism affects social services  | Study design: one-group pretest–posttest<br>Outcomes measured: cultural competence<br>Instruments used: study survey  | Assessment time points: pre-training; 2–7 years<br>Analysis: NR<br>Effect(s): null improvements   |
| 8. Yang <i>et al.</i> [44]    | N = 9 participants<br>Demographics: of initial 12 applicants, there were 7 White, 1 Black, 3 Hispanic<br>Setting: education   | Theory: Dialectical Behavior Therapy (DBT)<br>Modality: online<br>Duration: one 3-hour training over 2 days<br>Facilitator characteristics: NR<br>Topics covered: history of racism, impact of racism on students, antiracism, DBT skills   | Study design: one-group pretest–posttest<br>Outcomes measured: confidence, knowledge, ability to teach antiracism<br>Instruments used: study survey, Black-White Implicit Association Test (optional)   | Assessment time points: Pre and post-training, 6 months (no data reported)<br>Analysis: NR<br>Effect(s): null improvements in confidence, knowledge, ability to teach antiracism.   |

NR, not reported.  
<sup>a</sup> Statistically significant refers to  $P < 0.05$ .

short-term changes in *knowledge and awareness*, such as improvements in knowledge of specific terms [25], cultural knowledge [18], cultural essentialism [24], enhanced self-awareness [20], and perceived importance of recognizing biases and improving diversity [25, 27]. Four studies (26.7%) reported significant changes in participants' *attitudes*, such as increased participant self-efficacy (confidence) to act in accordance with, amplify, and teach DEI principles [29, 30], and improved positive feelings toward gender outgroups [21, 23]. Three studies (20%) reported significant differences in participants' *skills or behavioral changes*, including improved conflict resolution [26], increased familiarity with skills needed to work in culturally and racially diverse teams [18] and fostered inclusive environments [25], and being less likely to exhibit less racial prejudice [18].

### Antiracism training search results

The literature search for antiracism training studies yielded a total of 149 nonduplicate articles (Figure 2). Of those,  $N = 94$  articles were excluded based on the title review,  $N = 39$  articles were excluded during abstract review, and  $N = 8$  articles were excluded during full-text review, resulting in a total of eight articles eligible for inclusion. Table 2 presents a summary of characteristics across antiracism training studies included in this review.

### Antiracism training: Participant, setting, and training characteristics

The number of participants in antiracism training studies in this review varied from 9 to 875, and 6 of the 7 studies reported having at least 55% White participants. Of the five studies that reported gender demographics, four had participant samples that were 75% or more female participants. Studies of antiracism training took place in the following settings: secondary education ( $N = 3$  studies, 37.5%), health care ( $N = 3$  studies, 37.5%), social work ( $N = 1$  study, 12.5%), and community service providers ( $N = 1$  study, 12.5%). Trainings were conducted in-person ( $N = 6$  studies, 75%) or online ( $N = 2$  studies, 50%). No studies utilized a hybrid approach to delivering trainings.

Seven out of the eight antiracism training studies utilized a one-time training session. Antiracism trainings lasted 1 day on average and ranged from 3-hour to 3-day sessions. The trainings covered a variety of topics such as bias ( $N = 4$  studies, 50%), stereotypes ( $N = 2$  studies, 25%), and cultural competency ( $N = 2$  studies, 25%), among others. Five studies explicitly identified theories or frameworks as informing the antiracism training, including cultural competency training theory [45], Kern's six-Step Approach to Curriculum Development for Medical Education [46], dialectical behavioral therapy [47], the five-stage model of white racial identity development [48], cultural humility [49], and self-determination theory [34].

### Antiracism training study design and measures

All studies of antiracism trainings utilized a single group pre-test/posttest study design with participants self-selecting to participate. None utilized a randomized design or included control or comparison groups. Participants in all studies completed self-reported, self-administered surveys. Five out of the eight antiracism training studies used a nonvalidated study survey to assess outcomes, with the remaining three studies using the Race Implicit Association Test [50], the White

Racial Identity Attitudes Scale [51], and the Color-Blind Racial Attitudes Scale [52].

The most frequently measured outcome across studies was attitudes ( $N = 5$  studies, 62.5%), followed by knowledge ( $N = 4$  studies, 50%), behavior ( $N = 3$  studies, 37.5%), and cultural competence ( $N = 1$  study, 12.5%). Of the two studies that collected and reported follow-up data, one reported data at 6 months post-training [38] and one reported data from participants 2–7 years post-training [37].

### Antiracism training impact

Five out of seven of the antiracism training studies reported statistically significant results for at least one measured outcome, including improvements in *knowledge and awareness* [37–40, 42] or *attitudes* (e.g. decreased color-blind attitudes, more open attitudes about racial issues, privilege, and institutional discrimination, increased confidence in applying antiracism knowledge in the workplace and teaching antiracism awareness to others [42, 44]) following completion of trainings. Three studies reported significant changes in *skills*, such as improved professional development skills [38] and cultural competence in serving diverse communities [43]. Only one study reported significant differences in behaviors or actions in the workplace (e.g. becoming familiar with the customs of local Black communities, changing pediatric posters to be more welcoming to diverse clients) [43].

### Patterns across DEI and antiracism training studies

Of the DEI studies that used multi-session trainings, 85.7% reported significant improvements in one or more measured outcomes compared with 62.5% of DEI training studies that used one-time sessions. With respect to antiracism training studies, 100% of those that used multi-session trainings reported significant improvements in one or more measured outcomes, compared with 83.3 of studies that used one-session trainings. Of the 14 studies that were theory informed, 85.7% reported improvement in one or more measured outcomes compared to 66.6% of the nine studies that did not report being informed by a specific theory. Nine of the 23 studies (39.1%) had long-term follow-up timepoints (longer than one month), ranging from 6 weeks to 7 years. Of these studies, only six (66.7%) reported follow-up data and three (33.3%) reported continued improvement at follow-up. All four of the studies that used randomized designs showed significant improvements in at least one measured outcome [21, 23, 26, 53].

### Discussion

This study systematically reviewed the characteristics, measures, and outcomes of DEI and antiracism training studies published over the past two decades. Of the 23 studies included in this review, 15 articles (65.2%) utilized a one-time training session and 19 were delivered in-person (82.6%). Over 70% of studies had majority female participants. Sample sizes ranged widely from 5 to 3016 participants, with less than half ( $N = 9$  studies, 39.1%) having 100 or more participants. Training content, objectives, theoretical framework applied (if any), measures, and impact varied widely across studies. Few studies utilized randomized designs to test the training ( $N = 4$  studies, 17.4%) or collected follow-up data ( $N = 6$  studies, 26.1%). The majority of studies took place in education ( $N = 14$  studies, 60.9%) or health care ( $N = 4$

studies, 17.4%). Over half (60.9%) of studies in this review reported statistically significant changes in outcomes of interest, with improvements most often observed in participants' self-reported knowledge, attitudes, and awareness, and less frequently observed in participants' skills or behavior change. We noted certain trends, such as studies that used multi-session trainings, were theory informed, and utilized randomized designs had higher rates of yielding improvements in one or more outcomes; however, our sample size of studies was too small to make further analytic comparisons. Findings are consistent with prior research demonstrating wide variation in characteristics and effects of DEI and antiracism trainings [13–16, 54].

Results also highlight that the majority of research studies on DEI and antiracism trainings in this review are in the earlier phases of the translational research spectrum (T1: intervention development research phase and T2: efficacy research to determine intervention impact) [4, 5]. This review also included eight studies (five DEI and three antiracism trainings) published in the aftermath of the racial reckoning in the USA, a period when attention to community engagement and community-centered approaches was emphasized. Of these eight recent studies, only three mentioned community engagement to varying degrees. One study mentioned that facilitators who were knowledgeable about community engagement [25]; another described participants' interest in the community post-training [30]; and a third described a community partnership with the local school system and discussed the community need for antiracism training [44]. It is worth noting that due to the timeline of academic peer review and pandemic-related delays, these studies may have been initiated before the summer of 2020, and there may be relevant studies in progress or currently under review that have utilized community-centered approaches for the research.

Overall, few studies in this review regardless of publication date reported utilizing community-engaged or community-centered approaches, indicating that much opportunity exists to integrate communities into all aspects of the research, including research question conceptualization, intervention design and delivery, participant recruitment and retention, data collection measures and methods, data analysis and interpretation, dissemination of findings, and adapting and scaling interventions. Based on our review, we propose several strategies to advance the field of DEI and antiracism training through research and practice.

### Practice recommendation 1

Develop and deliver longitudinal trainings so that concepts and skills are able to be built upon over time. To enhance efficacy, both the frequency and duration of trainings can be extended beyond one-time sessions.

### Practice recommendation 2

Deepen the curriculum focus beyond individual knowledge and attitudes to include skill building and organizational development to support and sustain individual, interpersonal, and organizational change [55].

### Practice recommendation 3

Prioritize skills and behavioral and organizational change as outcomes of interest, using validated or established instruments where possible. While demonstrating improvements in

individual knowledge, awareness, and attitudes can be useful in assessing immediate gains from training participation, changes in these domains alone need to be accompanied by changes in skills and actual behavior change (individual, interpersonal, or organizational) in order to advance DEI and antiracism in the workplace.

### Practice recommendation 4

Use validated measures of DEI and antiracism where relevant. Studies in our review that reported using validated instruments in their outcome assessments include the Reaction-to-Diversity Inventory [56], Ambivalent Sexism Inventory [53], Race Implicit Association Test [34], and the Modern Racism Scale [57]. These four instruments have demonstrated reliability and validity in the studies included in our review as well as others [53,58–60]. Other validated instruments that practitioners and researchers can consider using include the Anti-Racism Behavioral Inventory [61], the Diversity Engagement Survey [62], and the Political Skill Inventory [63].

### Research recommendation 1

To enhance equity, integrate community-engaged approaches into the research design, implementation, evaluation, and dissemination. Examples include inviting community experts as co-leads or co-investigators on the research and including communities most impacted by discrimination to provide input and guidance.

### Research recommendation 2

To improve generalizability, future studies can recruit more representative samples (e.g. participants reflect organizational, sector, or regional demographics) and test trainings across a variety of sectors. The majority of studies in this review were conducted in higher education or health care, where female employees are overrepresented [64].

### Research recommendation 3

Studies of DEI and antiracism trainings can implement more rigorous study design methods to better identify best practices. Specifically, using a randomized controlled design as utilized by Ehrke *et al.*, Chang *et al.*, Holladay and Quinones, and Lindsey [21, 23, 26, 28] allows researchers and practitioners to better isolate the effect(s) of the training being tested, as this kind of design is the strongest for controlling for measured and unmeasured confounding. Additionally, utilizing comparison groups that receive some kind of DEI and/or antiracism training may be more ethical and equitable, so that all participants have some exposure to DEI and/or antiracist content (vs. an experimental group receiving training and a control group receiving no training or resources at all). For example, one study [21] randomly assigned participants to one of three experimental conditions (gender-bias training, general-bias training, or control). Another study used a  $2 \times 2$  design, where participants were randomly assigned to three conditions: training focus (similarities or differences), trainer gender (male or female), and trainer race (White or Black) [26]. Another study [28] used a  $3 \times 2$  design, where participants were randomly assigned to one of three diversity training methods (perspective taking, goal setting, or stereotype discrediting) and one of two groups of focus (African-Americans or the LGBT community). These methods can be implemented during the pilot phase of new training programs.

## Research recommendation 4

To measure change over time, studies of DEI and antiracism trainings can include intermediate and long-term follow-up assessments. As only six studies in this review (26.1%) collected and reported follow-up data, and of those, only one study focused assessing program satisfaction rather than intervention impact at the long-term follow-up [20], longer-term effects associated with such trainings remain not well understood. To address this gap, studies recommend an implementation of longitudinal design in diversity training interventions, which would improve data collection and allow researchers to ensure such interventions meet their goals and sustain improvements over time [65–67].

The above recommendations may facilitate the movement of DEI and antiracism training studies further along the research translational continuum. Study findings and recommendations should be interpreted within the context of the following limitations. The relatively small number of studies in our review limited our ability to conduct statistical comparisons of efficacy across study characteristics (e.g. multi-session vs. one-time training sessions, theory-informed vs. non-theory-informed interventions, randomized vs. non-randomized design). Of the studies that reported participant demographic data, several included majority female, majority White participants, which may limit the generalizability of findings to other populations. Additionally, studies did not include information on their overall workforce population; thus, participant representation and study findings may not necessarily be representative of the demographics of the organizations and industries from which the studies were conducted.

Study designs were moderate-to-weak with respect to causal inference. Combined with high variation in training content, duration, instruments, and outcomes assessed, it was not possible to determine which training characteristics are critical for improving outcomes. Studies also lacked uniformity in reporting the duration of training sessions (e.g. some only reported day-long sessions while others reported duration in hours) and in defining terminology. This study was limited to peer-reviewed research available through search engines and does not include white reports or publicly unavailable research. Depending on the organization, some recommendations may be more feasible than others. Lengthening the duration or frequency of training can represent a challenge in terms of time and resources. Organizations may have limited diversity among employees, thus limiting how representative participants may be with respect to demographics. More rigorous evaluation designs such as randomized conditions or comparison groups may be difficult to implement.

In conclusion, findings from this systematic review highlighted conceptual and methodological directions to enhance the delivery and efficacy of DEI and antiracism training interventions in the workplace. Areas for future research include exploring in-depth the intervention design elements that are critical to facilitate improved outcomes, such as examining the mechanisms and efficacy of different training durations, application of various theories and frameworks, and topics and skill sets addressed using strong evaluation designs and longer follow-up time periods. Future community-engaged research across the spectrum of translational research stages is needed to understand the optimal characteristics of DEI and antiracism trainings and how to best integrate these trainings within organizations for impact, sustainability, and scalability.

## Acknowledgements

M.L.W is supported in part by National Institutes of Health (NIH) National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Grant # R01DK120713 (PI: Wang). The study funder did not have any role over any of the following activities: study design, collection, analysis, or interpretation of data; writing the report; the decision to submit the report for publication. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the views of the NIH.

## Conflict of Interest Statement

The authors have no conflicts of interest to disclose.

## Author Contributions

M.W. conceived of the research questions and study design. A.G. and M.S. conducted the literature searches and synthesized data. All authors contributed to drafting and revising this manuscript and approved the final version.

## Ethical Approval

This article does not contain any studies with human participants performed by any of the authors. This article does not contain any studies with animals performed by any of the authors.

## Informed Consent

This study does not involve human participants and informed consent was therefore not required.

## Transparency Statements

This study was a systematic review of peer-reviewed studies; formal registration of the study and analysis plan was not applicable. There is no analytic code associated with this study. De-identified data is not applicable for this study. Methods and materials used to conduct this systematic review are stated in the manuscript.

## References

- McGregor J. Diversity job openings fell nearly 60% after the coronavirus. Then came the Black Lives Matter protests. *The Washington Post*. <https://www.washingtonpost.com/business/2020/07/15/diversity-jobs-coronavirus-george-floyd-protests/> (21 December 2022, date last accessed).
- O'Rourke J. Finance decision-makers outlook for 2021. *OneStream*. <https://onestreamsoftware.com/blog/finance-decision-makers-outlook-for-2021/> (20 December 2022, date last accessed).
- Ali Z. With global spending projected to reach \$15.4 billion by 2026, diversity, equity & inclusion takes the lead role in the creation of stronger businesses. *Cision*. <https://www.prnewswire.com/news-releases/with-global-spending-projected-to-reach-15-4-billion-by-2026--diversity-equity--inclusion-takes-the-lead-role-in-the-creation-of-stronger-businesses-301413808.html> (20 December 2022, date last accessed).
- Institute of Translational Health Sciences, University of Washington. *T-Phases of Translational Health Research*. <https://www.iths.org/investigators/definitions/translational-research/> (1 March 2023, date last accessed).

5. Knott CL. A journey with TBM: into the third chapter. *Transl Behav Med* 2022;12:723–5.
6. Bubenik S. Gender pay inequity: establishing fair and equal pay for women in your company. *Forbes*. <https://www.forbes.com/sites/scholleybubenik/2019/04/01/gender-pay-inequity-establishing-fair-and-equal-pay-for-women-in-your-company/?sh=8c087e58ed4b> (20 December 2022, date last accessed).
7. Gould E, Wilson V. *Black Workers Face Two of the Most Lethal Preexisting Conditions for Coronavirus—Racism and Economic Inequality*. Economic Policy Institute. <https://www.epi.org/publication/black-workers-covid/> (20 December 2022, date last accessed).
8. Day JC, Taylor D. *Do People With Disabilities Earn Equal Pay?* United States Census Bureau. <https://www.census.gov/library/stories/2019/03/do-people-with-disabilities-earn-equal-pay.html> (20 December 2022, date last accessed).
9. Dahl GB, Knepper MM. *Age Discrimination across the Business Cycle*. National Bureau of Economic Research No. 27581, 2020
10. Burns C. *The Gay and Transgender Wage Gap*. Center for American Progress. <https://www.americanprogress.org/article/the-gay-and-transgender-wage-gap/> (19 December 2022, date last accessed).
11. Lloyd C. *One in Four Black Workers Report Discrimination at Work*. Gallup. <https://news.gallup.com/poll/328394/one-four-black-workers-report-discrimination-work.aspx> (21 December 2022, date last accessed).
12. Sull D, Sull C, Zweig B. *Toxic Culture Is Driving the Great Resignation*. MIT Sloan Management Review. <https://sloanreview.mit.edu/article/toxic-culture-is-driving-the-great-resignation/> (3 January 2022, date last accessed).
13. Kalinoski ZT, Steele-Johnson D, Peyton EJ, et al. A meta-analytic evaluation of diversity training outcomes. *J Organ Behav* 2013;34:1076–104.
14. Paluck EL, Green DP. Prejudice reduction: what works? A review and assessment of research and practice. *Annu Rev Psychol* 2009;60:339–67.
15. Watt K, Abbott P, Reath J. Developing cultural competence in general practitioners: an integrative review of the literature. *BMC Fam Pract* 2016;17:158.
16. Renzaho AMN, Romios P, Crock C, et al. The effectiveness of cultural competence programs in ethnic minority patient-centered health care—a systematic review of the literature. *Int J Qual Health Care* 2013;25:261–9.
17. Emerick D. *What is Diversity Equity and Inclusion?* ESG. <https://www.esgthereport.com/what-is-diversity-equity-and-inclusion/> (3 January 2023, date last accessed).
18. Castillo LG, Brossart DF, Reyes CJ, et al. The influence of multicultural training on perceived multicultural counseling competencies and implicit racial prejudice. *J Multicult Couns Devel* 2007;35:243–55.
19. Cavaleros C, van Vuuren L, Delene V. The effectiveness of a diversity awareness training programme. *SA J Ind Psychol* 2002; 3:28.
20. Celik H, Abma TA, Klinge I, et al. Process evaluation of a diversity training program: The value of a mixed method strategy. *Eval Program Plan* 2012;35:54–65.
21. Chang EH, Milkman KL, Gromet DM, et al. The mixed effects of online diversity training. *Proc Natl Acad Sci USA* 2019;116:7778–83.
22. Davis DLF, Tran-Taylor D, Imbert E, et al. Start the way you want to finish: an intensive diversity, equity, inclusion orientation curriculum in undergraduate medical education. *J Med Educ Curric Dev* 2021;8:23821205211000352.
23. Ehrke F, Berthold A, Steffens MC. How diversity training can change attitudes: increasing perceived complexity of superordinate groups to improve intergroup relations. *J Exp Soc Psychol* 2014;53:193–206.
24. Fischer R. Cross-cultural training effects on cultural essentialism beliefs and cultural intelligence. *Int J Intercult Relat* 2011;35:767–75.
25. Harrison-Bernard LM, Augustus-Wallace AC, Souza-Smith FM, et al. Knowledge gains in a professional development workshop on diversity, equity, inclusion, and implicit bias in academia. *Adv Physiol Educ* 2020;44:286–94.
26. Holladay CL, Quinones MA. The influence of training focus and trainer characteristics on diversity training effectiveness. *Acad Manag Learn Educ* 2008;7:343–54.
27. Hostager TJ, De Meuse KP. The effects of a diversity learning experience on positive and negative diversity perceptions. *J Bus Psychol* 2008;23:127–39.
28. Lindsey AP, King E, Hebl M, et al. The impact of method, motivation, and empathy on diversity training effectiveness. *J Bus Psychol* 2015;30:605–19.
29. O'Connor MR, Barrington WE, Buchanan DT, et al. Short-term outcomes of a diversity, equity, and inclusion institute for nursing faculty. *J Nurs Educ* 2019;58:633–40.
30. Powell C, Yemane L, Brooks M, et al. Outcomes from a novel graduate medical education leadership program in advancing diversity, equity, and inclusion. *J Grad Med Educ* 2021;13:774–84.
31. Rawski SL, Conroy SA. Beyond demographic identities and motivation to learn: the effect of organizational identification on diversity training outcomes. *J Organ Behav* 2020;41:461–78.
32. Tillman F, Liu I, Lovince J, et al. Healthcare equity and leadership: implementation of diversity, equity, and inclusion training for pharmacy residents. *J Pharm Prac* 2022; 0:08971900221142684.
33. Tajfel H. *The Achievement of Inter-group Differentiation. Differentiation between Social Groups*. London: Academic Press, 1978.
34. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55:68–78. <https://doi.org/10.1037//0003-066x.55.1.68>
35. Flood R, Romm N. Diversity management: theory in action. *Syst Pract Action Res* 1995;8:469–82.
36. Mael FA, Tetrick LE. Identifying organizational identification. *Educ Psychol Meas* 1992;52:813–24.
37. Abramovitz M, Blitz LV. Moving toward racial equity: the undoing racism workshop and organizational change. *Race Soc Probl* 2015;7:97–110.
38. Johnson LM, Antle BF, Barbee AP. Addressing disproportionality and disparity in child welfare: evaluation of an anti-racism training for community service providers. *Child Youth Serv Rev* 2009;31:688–96.
39. Kingett J, Abrams D, Purewal K. *Show Racism the Red Card: Evaluating the Impact of Anti-racism Educational Interventions on the Attitudes of Young People in Secondary School Education..* Kent: University of Kent, 2017, 41.
40. Loe IM, Froehlich TE, Edrees HH, et al. Racism as an adverse childhood experience: an interactive workshop to train pediatricians to address racism in clinical care. *J Dev Behav Pediatr* 2021;42:502–511.
41. Martin P, Baxter A. *Mentoring African American Middle School Students: Applying Principles of Antiracism Education to the Problem of Closing the Black-White Achievement Gap*. Boston: University of Massachusetts Boston; 2000.
42. Santoro CM, Farmer M-C, Lobato G, et al. Antiracism training for nutrition professionals in the special supplemental nutrition program for women, infants, and children (WIC): a promising strategy to improve attitudes, awareness, and actions. *J Racial Ethn Health Disparities* 2022; 6:1–8.
43. Webb E, Sergison M. Evaluation of cultural competence and antiracism training in child health services. *Arch Dis Child* 2003;88:291.
44. Yang P, Crous Y, Balli-Borrero NA, et al. A ntiracism work in schools: using dialectical behavior therapy skills to empower South Texas educators. *J Am Acad Child Adolesc Psychiatry* 2022;61:1296–302.
45. Cabaniss ER, Frabutt JM, Kendrick MH, et al. Reducing disproportionate minority contact in the juvenile justice system: promising practices. *Aggress Violent Behav* 2007;12:393–401.
46. Thomas PA, Kern DE, Hughes MT, et al. *Curriculum Development for Medical Education: A Six-Step Approach*. Baltimore: Johns Hopkins University Press, 2015.

47. Linehan MM, Armstrong HE, Suarez A, *et al.* Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Arch Gen Psychiatry* 1991;48:1060–4.
48. Helms JE. *Black and White Racial Identity: Theory, Research, and Practice*. Westport, Conn: Praeger, 1993.
49. Tervalon M, Murray-Garcia J. Cultural humility versus cultural competence: a critical distinction in defining physician training outcomes in multicultural education. *J Health Care Poor Undeserved* 1998;9:117–25.
50. Greenwald AG, McGhee DE, Schwartz JLK. Measuring individual differences in implicit cognition: the implicit association test. *J Pers Soc Psychol* 1998;74:1464–80.
51. Helms JE, Carter RT. White racial identity attitude scale. *APA Psych-Tests* 1990.
52. Neville HA, Lilly RL, Lee RM, *et al.* Construction and initial validation of the color-blind racial attitudes scale (CoBRAS). *J Couns Psychol* 2000;47:59–70.
53. Salomon K, Bosson JK, El-Hout M, *et al.* The experiences with ambivalent sexism inventory (EASI). *Basic Appl Soc Psych* 2020;42:235–53. <https://doi.org/10.1080/01973533.2020.1747467>.
54. Hassen N, Lofters A, Michael S, *et al.* Implementing anti-racism interventions in healthcare settings: a scoping review. *Int J Environ Res Public Health* 2021;18:2993.
55. Valandra EC, Yazzie R. *Colorizing Restorative Justice: Voicing Our Realities*. St. Paul, Minnesota: Living Justice Press, 2020.
56. De Meuse KP, Hostager TJ. Developing an instrument for measuring attitudes toward and perceptions of workplace diversity: an initial report. *Hum Resour Dev Q* 2001;12:33–51. [https://doi.org/10.1002/1532-1096\(200101/02\)12:1<33::AID-HRD-Q4>3.0.CO;2-P](https://doi.org/10.1002/1532-1096(200101/02)12:1<33::AID-HRD-Q4>3.0.CO;2-P).
57. McConahay JB. Modern racism, ambivalence, and the Modern Racism Scale. In: *Prejudice, Discrimination, and Racism*. Cambridge, Mass: Academic Press, 1986, 91–125.
58. McConahay JB. Modern racism and modern discrimination: the effects of race, racial attitudes, and context on simulated hiring decisions. *Pers Soc Psychol Bull* 1983;9:551–8. <https://doi.org/10.1177/0146167283094004>.
59. Ivory KD, Dwyer P, Luscombe G. Reactions to diversity: using theater to teach medical students about cultural diversity. *J Med Educ Curric Dev* 2016;3. <https://doi.org/10.4137/jmecd.S37986>.
60. Greenwald AG, Poehlman TA, Uhlmann EL, *et al.* Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *J Pers Soc Psychol* 2009;97:17–41. <https://doi.org/10.1037/a0015575>.
61. Pieterse AL, Utsey SO, Miller MJ. Development and initial validation of the anti-racism behavioral inventory (ARBI). *Couns Psychol Q* 2016;29:356–81.
62. Person SD, Jordan CG, Allison JJ, *et al.* Measuring diversity and inclusion in academic medicine: the diversity engagement survey. *Acad Med* 2015;90:1675–83. <https://doi.org/10.1097/acm.0000000000000921>.
63. Ferris GR, Treadway DC, Kolodinsky RW, *et al.* Development and validation of the political skill inventory. *J Manage* 2005;31:126–52. <https://doi.org/10.1177/0149206304271386>.
64. U.S. Bureau of Labor Statistics. *Women in the Labor Force: A Databook*. U.S. Bureau of Labor Statistics. [https://www.bls.gov/opub/reports/womens-databook/2020/home.htm#:~:text=\(-See%20table%2012.\),and%20hospitality%20\(51.2%20percent\)](https://www.bls.gov/opub/reports/womens-databook/2020/home.htm#:~:text=(-See%20table%2012.),and%20hospitality%20(51.2%20percent)) (3 January 2023, date last accessed).
65. Elliott TC. How do we move the needle?: building a framework for diversity, equity, and inclusion within graduate medical education. *Fam Med* 2021;53:556–8. <https://doi.org/10.22454/FamMed.2021.199007>.
66. Devine PG, Ash TL. Diversity training goals, limitations, and promise: a review of the multidisciplinary literature. *Annu Rev Psychol* 2022;73:403–29. <https://doi.org/10.1146/annurev-psych-060221-122215>.
67. Avey JB, Luthans F, Mhatre KH. A call for longitudinal research in positive organizational behavior. *J Organ Behav* 2008;29:705–11. <https://doi.org/10.1002/job.517>.