Psychological Bulletin
2016, Vol. 142, No. 11, 1227–1274

© 2016 American Psychological Association
0033-2909/16/$12.00 http://dx.doi.org/10.1037/bul0000067

A Meta-Analytical Integration of Over 40 Years of Research on Diversity Training Evaluation

Katerina Bezrukova
University at Buffalo

Chester S. Spell
Rutgers University

Jamie L. Perry
Cornell University

Karen A. Jehn
University of Melbourne

This meta-analysis of 260 independent samples assessed the effects of diversity training on 4 training outcomes over time and across characteristics of training context, design, and participants. Models from the training literature and psychological theory on diversity were used to generate theory-driven predictions. The results revealed an overall effect size (Hedges g) of .38 with the largest effect being for reactions to training and cognitive learning; smaller effects were found for behavioral and attitudinal/affective learning. Whereas the effects of diversity training on reactions and attitudinal/affective learning decayed over time, training effects on cognitive learning remained stable and even increased in some cases. While many of the diversity training programs fell short in demonstrating effectiveness on some training characteristics, our analysis does reveal that successful diversity training occurs. The positive effects of diversity training were greater when training was complemented by other diversity initiatives, targeted to both awareness and skills development, and conducted over a significant period of time. The proportion of women in a training group was associated with more favorable reactions to diversity training. Implications for policy and directions for future research on diversity training are discussed.

Keywords: diversity training, diversity education, training effectiveness, bias, discrimination

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

During a diversity training session, a female participant was forced to stand up in front of her colleagues as an example of "the privileged White elite." Later in the same session, the consultant again asked her to stand proclaiming, "We all know who the most beautiful woman in the room is. It's the woman with the three private [school] degrees and the blond hair and the blue eyes." The woman remained in her seat, sobbing (MacDonald, 1993).

As this example illustrates, the effectiveness of some diversity training programs is open to question. Yet, the increasing demand for diversity training due to changing workforce demographics, globalization, continuing litigation, and other trends calls for a better understanding of the types of programs that can be effective. The case of a gay Rutgers University student's suicide after being harassed by fellow students (Schwartz, 2010) and similar incidents are a reminder that intimidation, discrimination, and threats to privacy based on individual differences remain commonplace. Other incidents over the past few years—the incident in Ferguson, Missouri, and the resulting civil unrest; the killings of New York police; and the massacre at the offices of the newspaper Charlie Hebdo in Paris, France—while different on many levels, all share one theme: They demonstrate how lives are at stake when differences between people are not accepted. As a response to this social issue, diversity training has the potential to make a huge, positive impact because diversity training strives to address prejudice, stereotyping and other biases (King, Gulick, & Avery, 2010).

Diversity scholars have indeed given generous attention to this topic, providing guidelines for successful implementation and evaluation of diversity training (e.g., Curtis, Dreachslin, & Sinioris, 2007; Ely, 2004; Kalev, Dobbin, & Kelly, 2006; Kalinoski et al., 2013; Kulik & Roberson, 2008a, 2008b; Paluck, 2006; Paluck & Green, 2009). This is clearly an area of psychology that can contribute to the enhancement of human well-being and society. The American Psychological Association (APA) has described diversity education as one of the five major learning goals for undergraduate education (American Psychological Association [APA], 2013). The variety and sheer numbers of training methods and designs employed on campuses and in organizations attests to how widespread the practice of diversity education has become.

This article was published Online First September 12, 2016.

Katerina Bezrukova, School of Management, University at Buffalo; Chester S. Spell, School of Business, Rutgers University; Jamie L. Perry, College of Business–School of Hotel Administration, Cornell University; Karen A. Jehn, Melbourne Business School, University of Melbourne.

This research was supported by the Santa Clara University Thomas Terry Grant. We thank Christy Arsky for her assistance with data coding. We are grateful to Michelle Campion and students from the "Managing Diverse Workforce" course at Santa Clara University for helping with additional data collection.

Correspondence concerning this article should be addressed to Katerina Bezrukova, School of Management, University at Buffalo, 274 Jacobs Management Center, Buffalo, NY 14260. E-mail: ybezruko@buffalo.edu

But what is the evidence that diversity training and education "works?" Reviews on diversity training and education conclude the evidence is decidedly mixed. At worst, diversity training has been shown to backfire in some cases by reinforcing stereotypes and prejudice among students (Robb & Doverspike, 2001) or creating new problems for the company (Kaplan, 2006), such as when air traffic controllers sued the Federal Aviation Administration because they had found diversity training traumatic (Epstein, 1994). Yet, other evidence suggests that diversity training can be effective. There are studies demonstrating how diversity training can reduce prejudice among students, enhance multicultural skills of nurses and medical students, improve productivity and engagement of diverse employees, and help retention of women and people of color in the workplace (cf. Anand & Winters, 2008; Rudman, Ashmore, & Gary, 2001; Smith, Constantine, Dunn, Dinehart, & Montoya, 2006). But little has been done to reconcile the opposing effects of diversity training to understand when and how diversity training promotes positive changes in trainees' learning outcomes.

Part of the problem is that researchers have approached the topic of diversity training evaluation using a wide array of theoretical interests, conceptualizations, and evaluations, both across and within disciplines (e.g., Bezrukova, Jehn, & Spell, 2012; Curtis et al., 2007; Kulik & Roberson, 2008a; Paluck & Green, 2009; Pendry, Driscoll, & Field, 2007; Smith et al., 2006). The volume of research on diversity training is remarkable, growing exponentially in the past decade, spanning a range of disciplines from psychology to education, social work, nursing and medical fields, business, and sociology. Yet, the increasing fragmentation of knowledge generated by researchers in various diversity training subfields calls for a multidisciplinary approach as well as for an updated, comprehensive, quantitative review of studies on the effects of diversity training to better inform the current research and practice of diversity training, instruction, and education.

Diversity Training: Theoretical Considerations

Consistent with prior work on diversity instruction (Avery & Thomas, 2004) and training evaluation (Kulik & Roberson, 2008a), we define diversity training as a distinct set of instructional programs *aimed* at facilitating positive intergroup interactions, reducing prejudice and discrimination, and enhancing the skills, knowledge, and motivation of participants to interact with diverse others (Pendry et al., 2007). Diversity training is generally seen as a separate branch within the training literature because it often elicits more emotionally charged responses than other types of training (e.g., job effectiveness training; Alderfer, 1992; Hanover & Cellar, 1998; Law, 1998; Rynes & Rosen, 1995). It is a set of educational activities offered by a university or an organization to its students or employees.

Research on diversity training is ideally placed at the interface between psychological theory on diversity and the organizational reality of training programs. On one hand, psychological theory on diversity has made important contributions to understanding one's personal attitudes and beliefs, and diversity overall using a range of theoretical perspectives and experimental methodologies (Paluck & Green, 2009; Pendry et al., 2007; Pettigrew & Tropp, 2006). Much of this research has focused on understanding the effects of training interventions aimed at altering stereotype content based on the contact hypothesis (Allport, 1954). A key assumption of the contact perspective is that providing opportunities for positive and cooperative contact between members of groups previously hostile to one another reduces prejudice. This literature has examined numerous laboratory interventions to raise awareness about prejudice or change group interactions (cf., Paluck & Green, 2009) and reported many cases of contact-based principles guiding successful diversity training programs. Yet, beyond training delivery, research has overlooked the role of context in effective diversity training programs.

In turn, organizational researchers have complemented this body of work by focusing on the context and delivery of diversity training, mostly drawing upon motivation and learning theories (cf. Aguinis & Kraiger, 2009; Baldwin & Ford, 1988; Salas & Cannon-Bowers, 2001). This literature stresses that context, or where the training is situated and how it is positioned and reinforced, is a particularly important correlate of motivation toward diversity training and outcomes (Colquitt, LePine, & Noe, 2000). Also, models of work-based learning (e.g., Avery & Thomas, 2004; Foldy & Creed, 1999; Raelin, 1997) and research on trainees' motivation to learn from diversity training programs (Wiethoff, 2004) suggested that the two dimensions fundamental for learning, knowledge and practice, work better when combined an acquired in concert (Raelin, 1997). One of the key principles in this area has been infusing training with "variety," or providing perspective-broadening information (Avery & Thomas, 2004) to best cater to all participants (Huber, 2013).

Our goal is to summarize the existing literature on diversity training and education by integrating psychological theory on diversity with organizational research on training models. To accomplish this goal, we build on Bezrukova and colleagues' (2012) narrative review and framework that differentiates between training inputs and outputs to delineate the constructs that should be included in our theory. According to this framework, training inputs include training context, design, and trainee characteristics, whereas outputs include participants' reactions to training and various learning outcomes. We add to this framework by providing a theoretical explanation for the main effects as well as boundary conditions behind the diversity training effects. Identifying the conditions (when something works or does not) is one of the main contributions of our article.

The Meta-Analysis

A significant part of our theoretical model centers on the outcomes of diversity training because the results of diversity training evaluation have been mixed. For instance, in a prior meta-analysis of 65 studies on diversity training, Kalinoski and colleagues (2013) found mixed evidence for attitude change (e.g., being more tolerable toward diversity as a result of the training), but larger effect sizes for skill-based outcomes. In contrast, a narrative review of 74 studies (Kulik & Roberson, 2008a) indicated consistently positive effects of diversity training on participants' overall attitudes and knowledge about diversity but less consistent effects on diversity-related skills and behavior. Hence, one of our objectives is to reconcile these contradictory findings by revisiting diversity training outcomes.

Turning to the moderators of diversity training effects, most work has been done in a rather piecemeal fashion identifying a

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

number of critical factors, but lacking an overarching theoretical framework. The few large-sample studies examining the impact of common diversity practices within or across organizations (Ely, 2004; Kalev et al., 2006) have been very instrumental in adding to our understanding of the effectiveness of various organizational efforts in promoting diversity. However, as Kulik and Roberson (2008b) pointed out, little research has focused on identifying which characteristics of diversity training are associated with positive outcomes (e.g., increased knowledge or changes in behaviors). Better understanding of effective design characteristics is needed to make sure our efforts in developing diversity training programs are not misplaced (Holladay, Knight, Paige, & Quinones, 2003). Thus, another objective of our study is to evaluate which design characteristics are associated with larger effect sizes.

We further build on prior research but, in contrast to past meta-analyses on diversity training (e.g., Kalinoski et al., 2013; Smith et al., 2006), we take a comprehensive, multidisciplinary perspective (see Table 1 for comparison across meta-analytical studies). We extend prior meta-analyses on diversity training (e.g., Kalinoski et al., 2013; Smith et al., 2006) in three critical ways: theoretical/practical contributions; data, scope, and types of analyses; and substantive conclusions. First, while Smith and colleagues' (2006) meta-analysis significantly contributed to the multicultural education literature and Kalinoski and colleagues (2013) added to our understanding of the effects of diversity training programs, neither provided a comprehensive theoretical framework. In response, we provide an integration of psychological theory on diversity with learning and motivation models from training literature to specify the context of training, design, trainees' characteristics, and methodological variables.

Second, we go beyond prior meta-analyses in terms of data, scope, and type of analyses considered. For example, unlike Kalinoski and colleagues' (2013) study, we crossed multiple search terms to obtain the broadest possible sample of relevant articles, and we provide a formal evaluation of publication bias in our sample, using a systematic approach that is transparent (Anderson et al., 2010; Kepes, Banks, McDaniel, & Whetzel, 2012). Third, in terms of substantive conclusions, we go beyond prior analyses by including participant reactions in addition to other outcome measures, and considering the theoretical underpinnings of the short- and long-term effects of diversity training. Our focus is on adult trainees, as, consistent with the sensitive period hypothesis (Uttal et al., 2013), relevant interventions for reducing prejudice among children and teens deserve their own review. Our primary audience is diversity training researchers, but educators, consultants, and managers alike may also benefit from knowing when and how to effectively implement diversity training.

Diversity Training Outcomes

Following the approach of Holladay and Quinones (2008), we examine the effects of diversity training on trainee cognitive, behavioral, and attitudinal/affective learning, as well as on trainee reactions (Kirkpatrick, 1959; Kirkpatrick & Kirkpatrick, 2006; Kraiger, Ford, & Salas, 1993). Cognitive learning refers to the extent to which trainees acquire knowledge (e.g., knowledge about cultural diversity issues). Behavioral learning concerns the development of trainees' skills, assessed via self-reports or implicitly identified skills (e.g., situational judgment tests based on a set of scenarios, Hauenstein, Findley, & McDonald, 2010), but also objective behaviors and results (e.g., performance evaluation of a trainee by a manager or trained observers, Juarez et al., 2006; Sanchez & Medkik, 2004). Attitudinal/affective learning captures changes in trainees' attitudes on diversity and self-efficacy (beliefs in their capacity to perform).

We propose that reactions to training will have stronger effects than all other learning outcomes. Reactions include trainees' feelings toward an instructor or trainer as well as toward the training overall (Holladay & Quinones, 2005; Rynes & Rosen, 1995). While studies comparing reaction to training and other outcomes are rare, Kirkpatrick's (1996) widely known training model suggests that participants do need to see their training as effective and worthwhile; training reaction is an antecedent of learning that leads to behavior (Giangreco, Carugati, Sebastiano, & Bella, 2010). As emotional responses (Kirkpatrick, 1979), we expect reaction effects to be more intense compared to other outcomes of diversity training.

> *Hypothesis 1:* Diversity training will have stronger effects on participants' reactions relative to cognitive, behavioral, and attitudinal learning.

Recent research on diversity training has not paid much attention to short- and long-term consequences of training. Thus, we examine both short- and long-term training effects. Short-term evaluation typically occurs at the end of the diversity training. For instance, participants complete rating scales at the end of the semester (in the case where a diversity course is being evaluated) after completion of their final exam (Murphy, Park, & Lonsdale, 2006). Long-term effects of diversity training mainly result from more permanent changes in beliefs, expectations, scripts, attitudes,

Table 1
A Comparison of the Sizes of Recent Meta-Analyses of Diversity Training Effects to That of the Current Meta-Analysis

Meta-analysis	Number of studies	*K*	*N*	Discipline	Theory development	Publication bias
Smith et al., 2006						
Meta-analysis 1 (retrospective)	45	45	5,991	One	No	Assessed
Meta-analysis 2 (prospective)	37	37	2,132	One	No	Assessed
Kalinoski et al., 2013	65	97	8,465	Various	Attitude theory and training	No Assessment
Present article	260	440	29,407	Various	Psychological theory on diversity and training models	Assessed

Note. K = number of effect sizes used in analysis; N = number of participants.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

and other factors. Long-term evaluation typically occurs after some time lag (ranges = from 1 month to 4 years, $M = 7.23$ months, $SD = 5.49$ in our data).

We predict that cognitive learning will persist, whereas attitudinal and behavioral learning will subside over time. We build on a prompting theory perspective (Sitzman, Ely, Brown, & Bauer, 2010) that suggests that participants use prompting from environmental stimuli to self-regulate and enhance learning after training (Carver & Scheir, 1990; Winne, 2005). Prompting can occur through cues in the workplace or while reading media stories related to diversity that reminds trainees of scenarios or situations (e.g., "Wow, this is what we've learned about Mexican culture in that diversity training half a year ago . . .") reinforcing cultural knowledge over time. Because cognitive based learning is about accumulation of verbal knowledge (Kraiger et al., 1993), it would be expected to increase in response to subsequent prompts; the effect over time would hence be stronger.

As for attitudes, or an individual's propensity to evaluate an entity as favorable or not (Eagly & Chaiken, 2007), changes in them can be triggered by training itself or prompts outside the training. These attitudes may gravitate back to the original evaluative judgments after the diversity training ends if attitudes that a person had before the training are reinforced (e.g., a media report or political speech that casts immigrants in a negative light may cause them to reevaluate how they interact with immigrants). Because attitudes are acquired behavioral dispositions (Eagly & Chaiken, 2007), similarly, environmental prompts can retard and even reverse skill development (Kraiger et al., 1993). For instance, individuals may start reconsidering behaviors if environmental cues cause them to question what they learned, resulting in a decline of behavioral learning outcomes over time.

Hypothesis 2: The effects of cognitive learning will persist, whereas the effects of attitudinal and behavioral learning will subside over time.

Theoretically Based Moderators

Diversity Training Context

The organizational literature on training transfer has repeatedly shown that motivations, expectations, and attitudes surrounding transfer are driven in part by context (Quinones, 1997; Tesluk, Farr, Mathieu, & Vance, 1995). Psychological theory on diversity has also made contributions: The "authority sanction" condition of the contact hypothesis suggested that contact among different people should be blessed by higher authority (Paluck, 2006), stressing the importance of institutional support as context. Following Bezrukova and colleagues (2012), in addition to differentiating between academic and organizational settings (Kulik & Roberson, 2008a), we also consider other contextual aspects such as the training approach (standalone vs. integrated with other diversity practices) and training attendance (mandatory vs. voluntary). Yet, we go beyond existing frameworks by developing theoretically driven predictions linking these three seemingly different contextual factors through a common mechanism—motivation to learn—to all outcomes of diversity training.

Settings: Organizational versus educational. The main goal of most diversity training programs in organizational settings is to create an inclusive workplace culture that is effective for relationships both inside and outside of an organization (Cox & Blake, 1991; Mor-Barak, 2005; National Urban League, 2009; Thomas, 1996). However, diversity training is often seen as being an "add-on" or secondary to the core purpose of the organization. In contrast, the main goal of campus-based diversity training is to learn about diversity and prejudice, and apply concepts through experiential learning and structured opportunities for intergroup contact (Avery & Thomas, 2004; King et al., 2010). Because goals drive motivation to learn (Locke & Latham, 1990), diversity training efforts in educational settings would be more aligned with overall mission of the school and thus could be more effective on campuses than in the workplace.

Approach: Standalone versus integrated. Diversity training can be done in one brief session (a "check-off-the-box" commitment), for example, covering legal and compliance issues or learning about cultural differences (Anand & Winters, 2008). Alternatively, diversity efforts can be part of a comprehensive diversity curriculum (Guy-Walls, 2007) or could be integrated and complemented by other diversity-related initiatives (Bendick, Egan, & Lofhjelm, 2001). This latter approach signals commitment and support for diversity from the top, and, according to the training motivation theory, can reinforce the motivation of trainees to learn (Salas & Cannon-Bowers, 2001). Bendick and colleagues' (2001) national sample of diversity programs and Rhyne's (1973) sample of junior high schools supported the view that integrated training was more effective than when it was isolated and standalone.

Attendance requirements: Mandatory or voluntary. We predict that mandatory training would be more effective as it sends a message that the organization is truly committed to diversity, thus increasing trainees' motivation to learn (Kellough & Naff, 2004; Paluck, 2006; Rynes & Rosen, 1994). This prediction is supported by training motivation theory (Colquitt et al., 2000) which stresses the importance of how the training is positioned and reinforced (Kraiger et al., 1993; Salas & Cannon-Bowers, 2001). The alternative of having a choice to attend diversity training may not lead to desired effects as documented in contact research (Pettigrew & Tropp, 2006), creating a situation when those attending are not the ones that would get the most out of the training (Ellis & Sonnenfeld, 1994). These arguments have received empirical support—mandatory attendance was positively associated with all learning outcomes of diversity training (D'Andrea, Daniels, & Heck, 1991).

Hypothesis 3: Diversity training will have stronger effects on all learning outcomes when the training context provides more motivation to learn (e.g., educational settings, integrated, and mandatory) than when it does not (e.g., organizational settings, standalone, and voluntary).

Diversity Training Design

Next, we focus on the design features of diversity training that prior research indicates may contribute to its effectiveness (Holladay & Quinones, 2008). Psychological models on intergroup conflict (cf., Paluck & Green, 2009) drawing on the contact hypothesis (Allport, 1954) have been particularly informative in theorizing about the effects of specific training design features. For example, personalized contact has been predicted to generate feel-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ings of familiarity between group members (Brewer & Miller, 1984), breaking down social boundaries while preserving recognition of group differences and improving cooperative behavior (Ensari & Miller, 2006). Ultimately, this provides insights into training focus (group specific vs. inclusive) and training duration. In turn, theory-based learning models from the training literature provide guidelines on designing diversity-specific training types and instructional methods.

Focus: Group specific versus inclusive. Training focus that is group specific can target either one group (e.g., focus on race) or multiple groups (e.g., race, gender, age, sexual orientation etc.). The group-specific approach, whether focusing on one or many groups, has been often criticized as leading to intergroup differentiation and attitude polarization (Stratton, Canales, Armas, & Miller, 2006). Training focus that is inclusive highlights contributions of all, thus providing more opportunities for cooperative contact and minimizing intergroup tensions (Ivancevich & Gilbert, 2000). As an illustration, a group-specific approach might raise the question "What is wrong with this outgroup?" (e.g., "women need to learn to be more assertive"), where an inclusive approach might frame the question as "What is wrong with this organization that treats outgroups worse than white male ingroups?" Inclusive-focused training deemphasizes group-specific issues to focus instead on the inclusiveness of the organization's culture, using individual groups' experiences as illustrations of a phenomenon.

Duration: Short versus long. In the studies we analyzed, the duration of training ranged from as short as 30 min (Govern, 1997; Hanover & Cellar, 1998) to as long as 4 years (Guy-Walls, 2007). Theoretically, the contact hypothesis suggests that duration is important because more time participants spend together leads to liking and makes intergroup encounters comfortable and feel "right" (Pettigrew, 1998). This argument has been supported, demonstrating that diversity training may be more useful when it is longer (e.g., Caffrey, Neander, Markle, & Stewart, 2005; D'Andrea et al., 1991; Griswold et al., 2006). Thus, we predict

> *Hypothesis 4:* Diversity training will have stronger effects on all learning outcomes when the design features provide more opportunities for cooperative contact (e.g., inclusive and longer) than when they do not (e.g., group specific and shorter).

Types: Awareness, behavior based, or combined. Diversity training types include awareness only, behavior only, or a combination of both components (Bezrukova et al., 2012). Awareness training focuses on getting participants to be more aware of their *own* and *other* cultural assumptions, values, and biases (Robinson & Bradley, 1997; Baba & Hebert, 2004). Skill-building (behavioral) training educates participants on monitoring *one's own* actions and appropriate responses to specific differences, such as identifying and overcoming interracial communication barriers. Learning theories suggest that participants can better understand their behavior (being aware of why they are doing what they are doing), when diversity training combines both awareness and behavioral components rather than when it focuses on only one (Raelin, 1997). We thus propose that the combined training will maximize learning opportunities and will be most effective overall.

Instruction: Many methods versus one method. Diversity training can use *many* different instructional methods or just *one*. Lecture-based diversity training (Lee, Anderson, & Hill, 2006), training based on video materials (Chrobot-Mason, 2004), or a simulation exercise (e.g., BaFá BaFá, Bush & Ingram, 2001; Jane Elliott's "blue-eyes/brown-eyes," Stewart, LaDuke, Bracht, Sweet, & Gamarel, 2003) are common examples of diversity training based on only one instructional method. However, according to Kolb and Kolb (2005), effective learning occurs when a learner "touches all the bases" and combines different learning styles (e.g., feeling, thinking, acting, or reflecting), thus maximizing learning opportunities. Therefore, we predict

> *Hypothesis 5:* Diversity training will have stronger effects on all learning outcomes when design features maximize learning opportunities (e.g., combined and multi-instructional) than when they do not (e.g., awareness or behavior based and one instructional method).

Trainee Characteristics

Because the composition of training groups has been shown to be an important determinant of diversity training effectiveness (Roberson, Kulik, & Pepper, 2003), we next focus on trainees' race, gender, and age as the most frequently examined or reported in the diversity training literature (e.g., Ely, 2004; Kulik, Pepper, Roberson, & Parker, 2007). We expect that diversity training will have stronger effects in groups with higher proportion of participants who are minority, women, and young. People of color and women tend to view themselves (or be viewed) as having higher levels of diversity acceptance (Smith et al., 2006), explaining some variation in diversity training's effects. Younger trainees may have an easier time with training content due to age-specific increases in certain components of memory, and faster cognitive response times (cf., Callahan, Kiker, & Cross, 2003). Therefore, we predict

> *Hypothesis 6:* Diversity training will have stronger effects on all learning outcomes when training groups consist of more women, minorities, and younger participants.

Study Rigor

Differences in research designs used in the evaluation of diversity training effects may reflect study rigor. Although survey (retrospective) studies form a substantial body of work in this literature, due to significant methodological weaknesses we do not include these studies in our review. Instead, we include observational or correlational studies based on a simple pretest- posttest design with no control group. We also include experimental studies based on the random assignment of participants to control or experimental groups. Finally, we include other designs with a control group (e.g., quasi-experiments with nonrandom assignments or posttests only). Our prediction here is based on contact theory that suggests that greater research rigor is routinely associated with larger effect sizes (Pettigrew & Tropp, 2006). This assertion makes sense when one takes into account the elimination of selection biases across experimental groups due to randomization of assignment to experimental conditions. Therefore, we predict

> *Hypothesis 7:* Study rigor will be associated with higher effect sizes.

Method

Literature Search

We conducted a literature search for published and unpublished studies on diversity training using online databases across multiple disciplines such as ABI-Inform, PsycINFO, PsycARTICLES, Business Source Premier, ProQuest Dissertations and Theses, Dissertation Abstracts International, COS Conference Papers Index, Social Services Abstracts, Sociological Abstracts, and the Education Resource Information Center (see Figure 1). We used multiple search terms to identify relevant diversity articles, such as *diverse*, *culture*, *multicultural*, *cross-cultural*, *pluralism*, *prejudice*, *bias*, *discrimination*, *sensitivity*, *tolerance*, *stereotype*, *race*, *racial*, *ethnic*, *ethnicity*, *Lesbian Gay Bisexual and Transgender (LGBT)*, *gay*, *age*, *older*, *generational*, *women*, and *gender*. We crossed these terms with the following training-related words: *train*, *workshop*, *education*, *course*, *curriculum*, *intervention*, *program*, *initiative*, *teach*, and *instruct*. In addition to the electronic databases, we hand-searched recent issues of psychology, organizational behavior, and human resource management journals to include articles that have not been published yet (in press) or available electronically. Next, the authors manually examined reference sections of published articles to help identify articles not included in the database searches. Finally, we searched the same databases using the names of researchers who had conducted research on diversity training.

Publication bias. Studies reporting large effects are more likely to be published than those reporting small or null effects (Rosenthal, 1979). We thus attempted to assess the effects of publication bias on our sample and analyses. First, we conducted a search of the Society for Personality and Social Psychology, Society for the Psychological Study of Social Issues, Society of Experimental Social Psychology, the Academy of Management, the Society for Industrial and Organizational Psychology, and APA conference programs through 2013 and identified 31 relevant articles. We then contacted authors of these articles and received 14 papers. Second, we conducted a search of ProQuest Dissertations and Theses and Dissertation Abstracts International, which

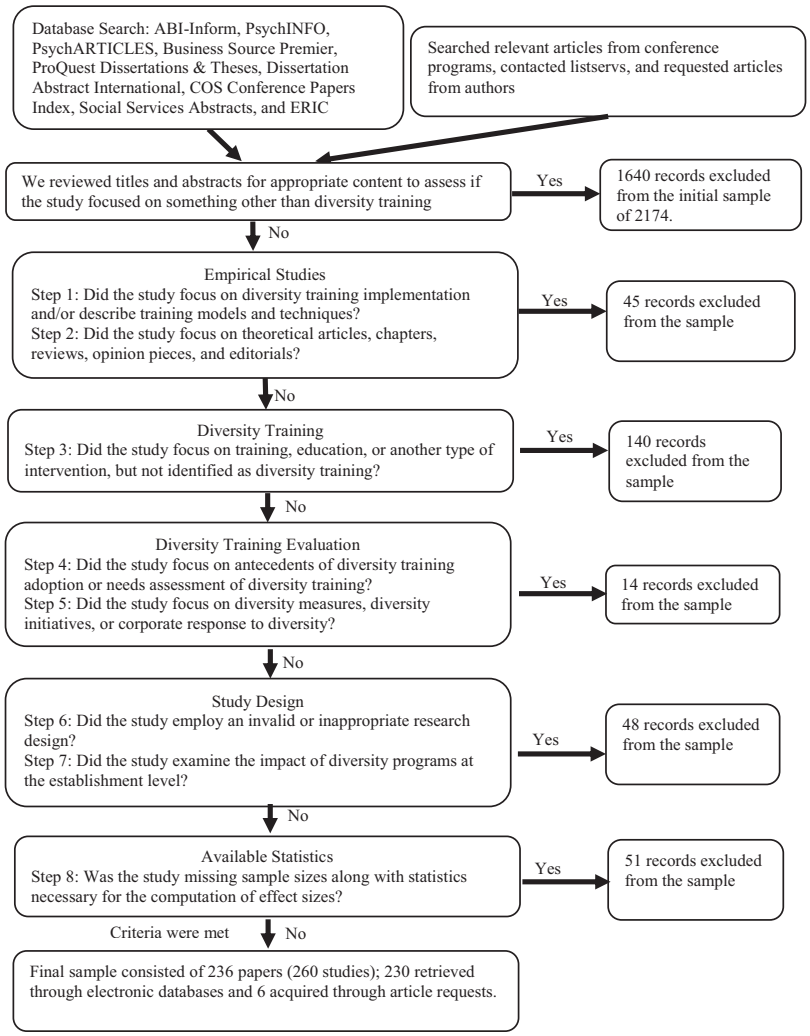

Figure 1. Flowchart illustrating the search and eliminating process for selecting articles.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

resulted in 764 papers (limiting the search to empirical studies and adult population). After an initial examination of titles and abstracts, 141 relevant dissertations were found. Finally, we posted requests for unpublished studies at six professional listservs (e.g., the American Sociological Association section on Social Psych listserv, Society for the Psychological Study of Social Issues listserve, hrdiv_net, OB Listserve, cmu-ingroup, gdo-l) and contacted authors and experts in research on diversity training asking explicitly for articles that might not yet have been published or not available electronically through 2013. Two researchers supplied studies that met the eligibility criteria. Altogether, unpublished studies accounted for 32% of included reports, a significant increase compared to past meta-analyses in this area.

Eligibility Criteria and Selection of Studies

The search procedure led to the identification of 2,174 articles that could potentially contain eligible studies of diversity training conducted on adults (1,086 published papers, 764 dissertations, 32 conference papers, and two papers from correspondence). This search was limited to adult populations and excluded all studies of diversity training, multicultural education, or any prejudice reduction intervention conducted on children, adolescents, or teens. After thorough examination of study abstracts, 534 articles were found relevant (359 published papers, 141 dissertations, 31 conference papers, and two papers from correspondence). The full text of these papers was obtained and reviewed. If a dissertation or a conference paper was eventually published, we augmented data from the published paper and counted the original dissertation and published article as one (published) study. Several criteria were further used to determine whether to include a study in the meta-analysis (see Figure 1 for steps in the selection of studies).

Empirical studies. Articles that provided recommendations/suggestions for successful diversity training implementation (e.g., Cox & Blake, 1991; Digh, 1998; Hollister, Day, & Jesaitis, 1993) and described different training models/techniques and approaches to diversity management (e.g., Gunsch, 1993; Hequet, 1991; Karp & Sammour, 2000; P. Remer & Remer, 2000) were not eligible, and thus were excluded (15 articles). Theoretical articles including chapters, reviews, and opinion papers or editorials were also excluded (30 records).

Diversity training. Consistent with other research on diversity training (e.g., Kulik & Roberson, 2008a, 2008b; Pendry et al., 2007), the study must satisfy two main criteria. First, it should be about *diversity*. The training intervention should aim at learning about other social groups, facilitating positive intergroup interactions, reducing prejudice and discrimination, and enhancing the skills, knowledge, and motivation of people to interact with diverse others (Pendry et al., 2007). Thus, studies that focused on training and development of disadvantaged social groups (36 records) were not eligible (e.g., Del Portillo, 1981; Gallman, 2006). Second, it should be about *training*. The study must have used training, education, or another type of intervention for teaching purposes and not exclusively for research. Hence, studies that focused on administering experimental manipulations to examine, for example, subjects' levels of prejudice (79 records) did not qualify (e.g., Anderson, 1982; Byrne & Wong, 1962; Kawakami, Dovidio, & van Kamp, 2007). Further, studies about diversity policies were not eligible, and this led to the exclusion of 25 records.

Diversity training evaluation. The study must have at least one diversity training outcome measure (e.g., reactions, cognitive, affective, or/and behavioral learning). Studies that examined the antecedents of diversity training adoption or provided needs assessment of diversity training were not eligible and hence were excluded from the dataset (six records). Studies that pointed out the importance of measuring diversity and creating a system of metrics to assess its impact (Thomas, 1999; Wheeler, 1998; Winterle, 1993), case studies that described diversity initiatives employed in specific organizations (Cornelius, Gooch, & Todd, 2000; Gilbert & Ivancevich, 2000; White, 1998), or explored the corporate response to the issue of diversity (Braham, 1989; Kleeb, 1989) were also not eligible and thus were excluded from the dataset (eight records).

Study design. The study must have employed a valid and appropriate research design, defined as meeting at least one of the following design criteria: (a) use of a pretest, posttest design that assessed training outcomes relative to a baseline measure obtained before the intervention was given; (b) inclusion of a control or comparison group, such as the comparison of training outcomes among employees who had versus did not have diversity training with or without the element of random assignment to treatment or control (McBurney & White, 2004). Studies that examined the impact of diversity programs at the establishment level (e.g., Kalev et al., 2006) or used retrospective survey methodology (e.g., Ely, 2004) were not eligible and were also excluded (48 records).

Available statistics and reporting. Sufficient information had to be present to calculate effect sizes. That is, studies had to report sample sizes along with the statistics necessary for the computation of effect sizes: group means and standard deviations; t, F, or chi-square values from between-groups analyses; precise p values and degrees of freedom from between-groups analyses; or other effect size values (e.g., correlation coefficient; Borenstein, Hedges, Higgins, & Rothstein, 2005; Lipsey & Wilson, 2001). When such data were not reported, the authors were contacted with a request for additional data. In these cases, five authors responded but none were able to provide additional data resulting in the exclusion of all 51 studies (see Figure 1). Overall, these selection criteria resulted in the inclusion of 236 articles, which comprised 260 eligible studies with 29,407 participants (coded 1,353 effect sizes).

Coding System and Decisions

All the authors and several research assistants coded studies that met the inclusion criteria. Research assistants received extensive training to ensure the reliability of their efforts. The coding system addressed the following characteristics of each study based on prior research: the publication status, the number of participants and their composition (% of women, % of white, average age and education), diversity training context in terms of setting (organizational vs. educational with employees in universities coded as organizational settings), approach (standalone vs. integrated), and attendance (required vs. voluntary), diversity training design in terms of focus (group specific vs. inclusive), training duration (in hours), type (awareness only, behavior only, and both awareness and behavior), and instruction (one vs. many), diversity training

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

outcomes (reactions, cognitive learning, attitudes/affective learning, and behavioral learning), the interval between the end of training and the collection of outcome measures (in months), study rigor (research design: observational/correlational, quasiexperimental, and experimental).

In developing our coding system, we followed the established guidelines for meta-analysis and coded each study based on the same criteria (Wilson, 2009; see Appendix A for the descriptions of included studies). Two of the authors independently coded 30% of the articles. The vast majority of the studies were coded similarly by both coders (mean overall Cohen's kappa = .914). Next, the two authors met to resolve coding discrepancies. They each discussed their rationale for how they coded the articles. After discussing their differences they agreed upon the best coding for the articles. All textual data were also read and analyzed by research assistants not associated with this project to control for potential bias in interpretation (Marshall & Rossman, 1989). No major discrepancies were found across the raters. This analysis confirmed previously identified themes and categories (e.g., context, design, trainee characteristics) in prior diversity training research (Bezrukova et al., 2012) and was also consistent with the more general training literature (e.g., Goldstein & Ford, 2002; Noe, 2010; Roberson et al., 2003; Werner & DeSimone, 2009).

To code the outcomes of diversity training, we have drawn on the Kirkpatrick (1959; Kirkpatrick & Kirkpatrick, 2006) and Kraiger and colleagues' (1993) models of training evaluation. Reactions were operationalized as self-report measures representing trainees' responses to the training program or trainer. Examples of reactions included scales that assess participants' backlash against the training and reactions toward the trainer or the program itself (e.g., Holladay & Quinones, 2008). Attitudinal/affective outcomes were operationalized as measures of internal states that derive attitudes (including self-efficacy). Examples of typical measures of attitudinal outcomes included self-assessments of attitudes toward ethnic groups (e.g., Europeans, Mexican American, African American; Tran, Young, & Di Lella, 1994), gender, social class, sexual orientation, language, ability and other differences (e.g., the Beliefs About Diversity Scale; Middleton, 2002) or intercultural issues (Klak & Martin, 2003). Other measures of attitudes included associations or semantic differential scales that assess attitudes implicitly (Cleveland, Festa, & Montgomery, 1988; Greenwald & Banaji, 1995; Greenwald, McGhee, & Schwartz, 1998). One example included an assessment of participants' attitudes toward older workers measured using this scale: 1 = *active* and 7 = *passive*, 1 = *productive* and 7 = *unproductive*, and so forth (Kulik, Perry, & Bourhis, 2000).

Cognitive learning outcomes were operationalized as measures of primarily declarative knowledge about diversity issues (Kraiger et al., 1993). Examples of cognitive outcomes included the knowledge subscale of the Multicultural Counseling Inventory (e.g., Castillo, Brossart, Reyes, Conoley, & Phoummarath, 2007; Williams, 2005) or Multicultural Awareness Questionnaire that assesses participants' knowledge about cultural diversity issues by calculating the total number of correct responses (Law, 1998). Finally, behavioral learning outcomes were operationalized as measures of behavior or behavioral intentions. Examples included assessments of trainees' abilities to resolve conflict (Holladay & Quinones, 2008) or exhibit behaviors believed to contribute to effective diversity management in the workplace (e.g., "openly discourages comments or jokes that perpetuate stereotypes or prejudice"; Hanover & Cellar, 1998). Other behavioral outcomes included content-analyzed behaviors or situational judgment tests (Hauenstein et al., 2010; Roberson, Kulik, & Pepper, 2001; Stroup, 1998), and performance evaluation of a trainee by a manager or trained observers (Juarez et al., 2006; Sanchez & Medkik, 2004). Following a similar procedure as described above, two of the authors independently coded 30% of the articles; interrater reliability was acceptable for all outcome variables (Cohen's kappa > .700). The remaining articles were split between the authors who completed the coding.

Computation and Analysis of Effect Sizes

In our choice of the effect size estimate, we have followed tradition in the training literature (Arthur, Bennett, Edens, & Bell, 2003; Keith & Frese, 2008; Taylor, Russ-Eft, & Chan, 2005; Uttal et al., 2013) and used Hedges' g (Hedges & Olkin, 1985; Rosenthal, 1991). This estimate is similar to a traditional *d*, but more precise as it includes a correction for biases due to sample size (Lench, Flores, & Bench, 2011; Uttal et al., 2013). The conventions typically used to interpret Cohen's *d* can be applied to Hedges' *g*: an effect size of .2 is considered small, .5 is considered moderate, and .8 is considered large (Cohen, 1988). We coded effect size direction uniformly; positive values indicated improved outcomes as a function of training (e.g., more acceptances of differences across people, etc.) and negative values indicated poorer outcomes as a function of training. Because studies with larger samples provide a more precise estimate of the effect size of interest, effect sizes were calculated with the small-sample correction formula for unbiased effect sizes by weighting the effect size associated with studies by same size (Hedges & Olkin, 1985).

We calculated an overall effect size for each study (see Appendix A). For the studies where outcome was measured both before (pretest) and after (posttest) training in both an experimental and control group, we calculated the overall effect size as the difference between the improvement in the experimental group and the improvement in the control group. For the within-subjects-only design, in which there was no control group and an outcome was measured before and after training, we calculated the effect size as the difference between the posttest and pretest. For between-subjects designs, we calculated the variance of effect sizes following Hedges and Olkin's (1985) procedures. For within-subjects designs, we calculated the variance of effect sizes following Morris (2000) and Morris and DeShon (2002) procedures. More specifically, we corrected for dependence between means by including correlations between dependent means in effect size calculations for these studies (Morris & DeShon, 2002). If correlations were not reported, we used available information to estimate them (e.g., *M*s, *SD*s, *t* statistics, see Morris & DeShon, 2002). Finally, when data were classified by subgroups (e.g., gender, race), we used the summary data from the subgroups to recreate the data for the study to compute the effect size and variance (Borenstein, Hedges, Higgins, & Rothstein, 2009, pp. 221–222).

Because effect sizes are usually computed from means and standard deviations, when these were not directly reported, we

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

transformed data reported in other formats (e.g., chi-square, correlation, etc.) to *g* coefficients. In 12 cases, no statistic was available but an analysis was reported as statistically significant, so we determined the *g* value corresponding to the reported alpha level. If the effect was reported as nonsignificant but no specific information was provided, effect sizes were estimated by presuming $p = .50$ (e.g., Lench et al., 2011). We estimated the average effect sizes for diversity training with outcome measures using a random-effects model. Comprehensive Meta-Analysis, Version 2 (Borenstein et al., 2005) was used for all our analyses as it provides a well-organized and efficient format for conducting and analyzing meta-analytic data (Uttal et al., 2013).

Independence of effect sizes. When a study reported multiple measures of a focal construct (e.g., two forms of written tests evaluating trainee's knowledge about cultural differences), we followed the recommendation by Geyskens, Krishnan, Steenkamp, and Cunha (2009) and created a single composite variable (obtained through averaging) for that construct to avoid violating the independence assumption by including multiple correlations from the same study (all correlations among the indicators were high). This consequently reduced the sample of effect sizes ($k = 440$). Because training outcome was a variable of interest, if a study reported effect sizes for multiple outcome types (e.g., cognitions, behaviors), these effects were considered to be independent, and thus were retained as separate data points. Similarly, because time was one of the moderator variables of interest, data points based on temporally repeated measures of the same or similar outcome variable for the same sample were considered to be independent and retained as separate data points (see Appendix B).

Moderator Analyses

We then assessed the amount of estimated variation in the effect size distribution using several indicators of heterogeneity (T, T^2, and I^2) and homogeneity (Q) of variance. A significant Q indicates the likelihood of moderators that explain variability in effect sizes across studies (Lipsey & Wilson, 2001). The I^2 statistic quantifies the degree of heterogeneity by estimating the percentage of the variance that is attributable to between-studies variability with percentages of $I^2 = 25$, 50, and 75 indicating low, moderate, and high degrees of heterogeneity respectively (Higgins & Thompson, 2002). To examine the effects of categorical moderators, we used a mixed-effects meta-analytic categorical test, the meta-analytic equivalent of analysis of variance. We used the between-groups goodness-of-fit statistic Q_B to test whether the categorical moderator model was statistically significant and then examined each subgroup within the sample by testing the confidence intervals for statistical significance and by comparing the effect sizes across subgroups. When significant, Q_B indicates that the effect sizes differ among the levels of the categorical moderator.

To examine the effects of continuous moderators, such as the length of time between the end of diversity training and collection of outcome measures, we used the metaregression program within Comprehensive Meta-Analysis with a maximum likelihood estimation procedure (Borenstein et al., 2005). Two indexes assessing the overall fit of the weighted regression model can be calculated: a Q_R attributable to the regression and a Q_E error or residual (both are distributed as a chi-square). Q_R is analogous to an F for a regression model and if significant, indicates that the regression model explains significant variability in effect sizes. Following recommendations of Arthur, Bennett, and Huffcutt (2001) we limited our moderation analysis to factors with two or more data points. Although there is no universally accepted cutoff as to the minimum number of studies to include in a meta-analysis, we acknowledge that using such a small number raises the possibility of second-order sampling error and concerns about the stability and interpretability of the obtained meta-analytic estimates (Arthur et al., 2001). However, we chose to use such a cutoff for the sake of completeness but emphasize that meta-analytic effect sizes based on less than five data points should be interpreted with caution.

Results

Study Characteristics

Analyses included 260 studies reported in 236 articles, 29,407 participants, and 1,353 effect sizes coded and aggregated to 440 for the main analysis (see above). Overall, 176 studies (68%) were published in journals and 84 studies (32%) were from (unpublished) dissertations (79 studies) or conference articles (five studies). Twenty-nine studies (11%, 3,837 participants) were conducted outside of the United States with the majority of studies from Canada (seven studies), Australia (five studies), United Kingdom (four studies), Spain (two studies), and South Africa (two studies). Study samples averaged more than 50% women. Participants' age varied from 18 to 75 years old. Most were White (73%) and from the United States (89%). The trainees' level of education varied from no college (eight studies) to advanced degree (12 studies) with the majority of participants having some college education (139 studies) or graduate degree (85 studies). The majority of trainees came from educational settings (23,429 participants) and were students (198 studies). The remaining 5,978 participants represented health professionals (30 studies), educators and scientists (14 studies), managers and nonmanagerial personnel (12 studies), public and social service workers (12 studies), and others (e.g., inmates, military, taxi drivers, and Rwandan genocide survivors). The time between the end of diversity training and outcome measures ranged from immediately after training to 24 months ($x = .83$, $SD = 3.09$). The median length of training was 28.04 hr ($SD = 37.91$).

Robustness Check and Sensitivity Analyses

Consistent with the Meta-Analysis Reporting Standards of APA and also in line with Kepes, McDaniel, Brannick, and Banks (2013) recommendations, we assessed the validity of our results with sensitivity analyses. The purpose of these analyses is to determine whether different decisions and assumptions made during the review process have substantially influenced the obtained results. As part of our sensitivity analyses, we assessed the potential causes of nonrobustness in terms of outliers, missing data, and publication bias.

Outliers. To detect outliers, we computed the sample-adjusted meta-analytic deviancy statistic (Huffcut & Arthur, 1995). We used a sample-adjusted meta-analytic deviancy cutoff of >4 (e.g.,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Steel & Kammeyer-Mueller, 2002) to determine outliers. On the basis of these analyses, one potential outlier was identified (Kitchens-Stephens, 2005). Sensitivity analyses were further conducted to assess the robustness of findings when including versus excluding this outlier. While inclusion of this outlier could distort the main analysis because it estimates a different population mean than the mean estimated by the remaining effect sizes (Lipsey & Wilson, 2001), the results did not differ significantly when the outlier was included.

Treatment of missing data. Sensitivity analyses were conducted to assess for potential biasing effects of missing data on the effect size estimate. That is, a conservative adjusted effect size estimate was computed by imputing an effect size of $g = 0.00$ for studies that were excluded solely because of missing data (Lipsey & Wilson, 2001). Then, the analysis with and without imputations was compared. An alternative imputation approach was used to assess the robustness of the meta-analytic results by substituting the mean of the nonmissing observations for the missing observations (Kepes et al., 2013). The results did not differ significantly across three types of analysis, thus providing confidence in our findings.

Assessing publication bias. We evaluated whether publication bias affected our results in several ways. First, we created a funnel plot of each study's mean weighted effect size versus its corresponding standard error to provide a visual measure of publication bias (see Figure 2). The contour-enhanced funnel plot had a symmetrical "funnel" shape suggesting no publication bias in our meta-analysis. To assess the potential impact of the missing studies on the overall effect size, we used Duval and Tweedie's (2000) trim-and-fill technique to impute the effect sizes associated with the potential missing studies. In this analysis, we estimated that there are 75 missing studies. When incorporating these studies, the effect size decreased from .38 to .20 (95% confidence interval [CI] [0.14, 0.23]). However, the value was still significant ($p < .001$), suggesting that the reported mean effect is not simply an artifact of publication bias. Egger's regression analysis resulted in a significant intercept: *intercept* = 1.51, $t(260) = 7.05$, $p = .00$. Yet, this test, particularly with sufficient statistical power, may detect potentially "trivial" bias (i.e., bias that has little impact on the conclusions; Kepes et al., 2012, p. 634).

As an additional measure of publication bias we examined a forest plot for evidence of "drift" in the cumulative point estimate (Borenstein et al., 2009). This process included performing a cumulative meta-analysis where effect sizes were added one at a time to the analysis and the mean effect size was recalculated each time until all the effect sizes have been added. No "drift" was evident in the forest plot where effect sizes were sorted by precision (the most precise effect size, i.e., the effect size from the largest sample, is added first, followed by the second most precise, and so on). Finally, because "the only true test for publication bias is to compare effects in the published studies formally with effects in the unpublished studies (Borenstein et al., 2009, p. 280), we directly compared the average effect size of published studies ($g = .36$, $SE = .03$, $k = 176$) and unpublished studies ($g = .41$, $SE = .05$, $k = 84$) in our sample, where k represents the number of studies. The difference was not significant ($p = .25$). We also checked on publication bias by rerunning all our analyses without the unpublished studies and found results were consistent across both sets of analysis. Taking together, these results suggest that while there was some evidence of publication bias in our sample, it seems unlikely that the major results are driven by publication bias. Also, we believe that the concern over publication bias can be further reduced because studies reported significant effects in one direction for some study variables and not for others and we included all comparisons. For example, integrated diversity training can be associated with better training outcomes, yet standalone diversity training may have limited effectiveness.

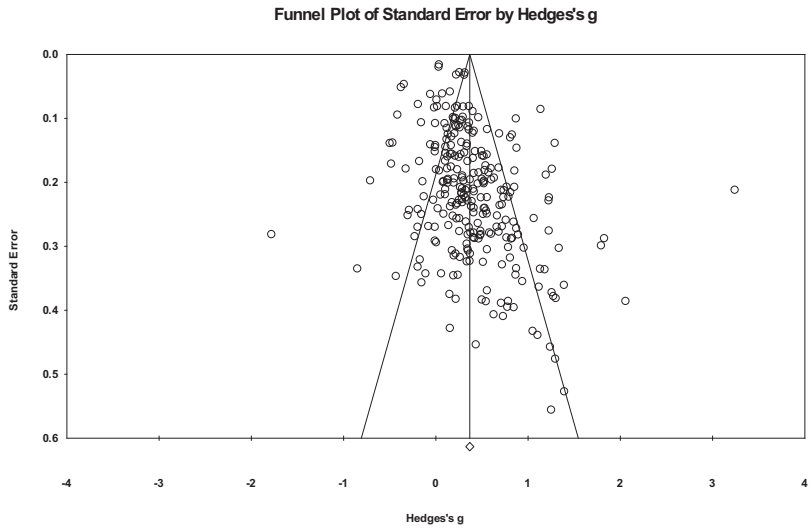

Figure 2. Funnel plot of publication bias: Standard Error × Hedges' g. The funnel plot presents the standard error on the *y*-axis and the Hedges' g effect size on the *x*-axis. White circles represent observed data points. On the *x*-axis, the white diamond represents the mean effect size.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Main Effects

The overall mean effect size across all studies included in the meta-analysis was significant ($g = .38$), which suggests that, in general, diversity training is associated with better training outcomes (see Table 2). The indicators of heterogeneity of the effect size distribution suggest that there is variation in the true effect size distribution ($T^2 = .10$, $T = .31$) and that a substantial amount of this variation can be explained by between-study differences ($I^2 = 85.73\%$). The overall effect size distribution contained more variation than would be expected by chance, $Q(259) = 1{,}828$, $p < .001$, suggesting that moderators might account for some of the variance in the effects. Next, we assessed whether the effectiveness of training varied systematically as a function of a specific training output (Hypothesis 1). Diversity training had the largest effect on reactions ($g = .61$), followed by cognitive learning ($g = .57$), behavioral learning ($g = .48$), and attitudinal/affective learning ($g = .30$), $Q_B(3) = 41.48$, $p = .00$.

Further, we were interested in determining whether different training outcomes are maintained over time (Hypothesis 2). We thus considered the relationship between study effect sizes and the length of time between the end of training and when posttests were administered. To accomplish this, we regressed study training effect sizes on the number of months between the end of training and when posttest was taken (ranged from 0 month to 24 months). For studies that included both immediate and delayed posttests, we, in line with Taylor and colleagues (2005), used the delayed posttest effect size. Overall, we found negative and significant relationships between training effects and the length of training-posttest interval for all learning outcomes except cognitive learning (effects were significant at .1 alpha level for behavioral learning). Taken together, these

Table 2
Summary of Associated Effect Size Statistics for Categorical Moderators

Characteristic	Between-class effect *(Qb)*	*k* (studies)	*N* (sample sizes)	Effect size (*g*)	95% CI
Overall effect size[a]		260	29,407	.38	[.33, .42]
Training context					
Setting	4.79†				
Organizational		63	5,978	.42	[.32, .53]
Educational		198	23,429	.36	[.31, .41]
Approach	5.27*				
Stand-alone		229	25,974	.36	[.32, .41]
Integrated		30	2,898	.57	[.40, .73]
Attendance	.52				
Mandatory		51	7,861	.42	[.31, .53]
Voluntary		198	20,108	.37	[.32, .43]
Training design					
Focus	.15				
One group specific		190	20,807	.39	[.33, .45]
Multiple group specific		27	3,449	.37	[.25, .48]
Inclusive		32	4,726	.37	[.26, .48]
Type	10.08*				
Awareness only		121	17,400	.31	[.24, .37]
Behavior only		11	816	.46	[.26, .66]
Awareness and behavioral		118	10,201	.46	[.38, .52]
Instruction	2.85				
Many		169	18,496	.38	[.31, .43]
One		79	10,725	.37	[.30, .45]
Methodological moderators					
Study rigor evaluation	3.05				
Experimental		63	6,408	.44	[.31, .56]
Quasi-experimental		100	12,739	.41	[.32, .49]
Observational		95	10,260	.33	[.27, .40]
Outcomes	41.48***				
Reactions		12	1,436	.61	[.35, .87]
Attitudes/Affective learning		229	26,448	.30	[.26, .35]
Behavioral learning		93	8,125	.48	[.40, .54]
Cognitive learning		104	8,414	.57	[.48, .66]

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; *k* = the number of studies providing information included in the analysis, where the mean across outcomes is used for studies with multiple outcomes (all of the selected outcomes, assuming independence, are used in reporting the effect sizes by "outcomes"); *N* = sum of the sample sizes of studies providing information included in the analysis; *g* = Hedges' *g*; CI = confidence interval.
[a] Please note that because four outcomes are argued to be conceptually distinct and focus on different facets of the criterion space, there is some question about the appropriateness of an "overall effect" size. Accordingly, this information is presented here for the sake of completeness.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3
Summary of Regression Results for Continuous Moderators

Training outcome	*N* of studies	*b*
Overall: Long-term durability of training effects (in months)[a]	258	−.015 (.009)[†]
Cognitive learning	101	.030 (.008)[**]
Attitudes/Affective learning	225	−.009 (.004)[*]
Behavioral learning	93	−.015 (.009)[†]
Reactions	12	−.049 (.016)[**]
Overall: The length of diversity training (in hours)[b]	215	.002 (.001)[**]
Cognitive learning	89	.003 (.001)[***]
Attitudes/Affective learning	196	.002 (.001)[***]
Behavioral learning	76	.003 (.001)[*]
Reactions	12	.021 (.004)[***]
Overall: trainee age[c]	138	.001 (.004)
Cognitive learning	53	.012 (.010)
Attitudes/Affective learning	121	−.003 (.004)
Behavioral learning	49	.003 (.007)
Reactions	8	−.009 (.011)
Overall: Trainee gender[d]	207	.087 (.111)
Cognitive learning	80	.133 (.306)
Attitudes/Affective learning	183	−.029 (.124)
Behavioral learning	73	−.079 (.195)
Reactions	11	1.861 (.413)[***]
Overall: trainee race[e]	170	.119 (.111)
Cognitive learning	62	.125 (.308)
Attitudes/Affective learning	150	178 (.131)
Behavioral learning	55	−.104 (.193)
Reactions	10	.960 (.738)

[a] Effect sizes were regressed on the time when the posttest was taken (ranging from 0 months to 24 months), using unrestricted maximum likelihood metaregression. A positive beta value indicates that larger effect sizes are associated with longer timeframe. [b] Effect sizes were regressed on the number of training hours (ranging from .33 to 320 hr). A positive beta value indicates that larger effect sizes are associated with longer diversity training. [c] Effect sizes were regressed on the mean age of trainees within the training group (ranging from 18 to 75 years old). A positive beta value indicates that larger effect sizes are associated with older participants, while a negative beta value indicates that larger effect sizes are associated with younger participants. [d] Effect sizes were regressed on the proportion of women participants (ranging from 0 [only male trainees] to 1 [all women]). A positive beta value indicates that larger effect sizes are associated with the higher proportion of women participants. [e] Effect sizes were regressed on the proportion of white participants (ranging from 0 (all minority training groups) to 1 (all White trainees). A positive beta value indicates that larger effect sizes are associated with the higher proportion of white participants.
[†] $p < .10$. [*] $p < .05$. [**] $p < .01$. [***] $p < .001$.

results suggest that, whereas reactions to training and attitudinal/affective learning appear to decay after training, cognitive knowledge (e.g., about different cultures) in contrast, is maintained over time after diversity training (see Table 3).

Moderator Analyses

Diversity training context (Hypothesis 3).

Settings: Organizational versus educational. We first analyzed differences in overall effect size as a function of diversity training settings (see Table 2). While a larger effect size was found for programs conducted in organizational settings ($g = .42$) versus those in educational settings ($g = .36$), this difference did not reach statistical significance, $Q_B(1) = 4.79$, $p = .09$. We further analyzed these data broken down by the type of outcome measured (see Table 4). As expected, studies conducted in educational settings had larger effect sizes for reactions ($g = .80$) than those in organizational settings ($g = .28$), $Q_B(1) = 6.43$, $p = .02$, thus suggesting that participants might enjoy diversity training more in educational settings than in organizations. No other comparisons were found significant (see Table 4).

Approach: Standalone versus integrated. We analyzed differences in overall effect size as a function of diversity training approach (see Table 2). As expected, studies that utilized an integrated approach reported higher overall effect sizes ($g = .57$) than those that utilized a standalone approach ($g = .36$), $Q_B(1) = 5.27$, $p = .02$. We further analyzed these data broken down by the type of outcome measured (see Table 5). Studies that utilized an integrated approach had larger effect sizes for attitudinal/affective ($g = .47$) and behavioral learning (g = .86) than studies that utilized a standalone approach ($g = .27$, $g = .42$, respectively); these differences were significant, $Q_B(1) = 7.15$, $p = .01$; $Q_B(1) = 5.11$, $p = .02$, respectively. While the results for cognitive learning were in the same direction and consistent with others, the difference in the effect sizes between integrated and standalone approaches was not significant for this outcome, $Q_B(1) = .53$, $p = .47$.

Attendance requirements: Mandatory or voluntary. We found no statistically significant difference in overall effect sizes across mandatory versus voluntary diversity trainings ($g = .42$, .37, respectively), $Q_B(1) = .52$, $p = .47$. We further analyzed these data broken down by the type of outcome measured (see Table 6). Studies that described mandatory diversity training had larger effect sizes for behavioral learning ($g = .63$) than studies that described voluntary training ($g = .42$). The effects were in a similar direction for cognitive learning, yet did not reach statistical significance, $Q_B(1) = 3.73$, $p = .09$. The effect was opposite for the reactions criteria (mandatory: $g = .37$; and voluntary: $g = .71$), $Q_B(1) = 7.49$, $p = .02$. No significant differences between effect sizes for attendance requirements were observed for attitudinal/affective learning (mandatory: $g = .36$, and voluntary: $g = .28$), $Q_B(1) = 1.66$, $p = .40$.

Diversity training design (Hypotheses 4 and 5).

Focus: Group specific versus inclusive. We first analyzed differences in overall effect size as a function of diversity training focus and found no significant differences in effect sizes across inclusive, one or multiple group-specific diversity training, $Q_B(2) = .15$, $p = .93$ (see Table 2). We further analyzed these data broken down by the type of outcome measured and once again found no statistically significant differences in effect sizes across all three training foci on all outcomes (see Table 7).

The length of diversity training. We explored whether the length of diversity training would be associated with an overall effect size as well as various outcome measures separately by regressing respective study effect sizes on the number of hours of diversity training. Table 3 presents unstandardized *b* coefficients representing the relationship between training effect sizes and hours of training for an overall effect size as well as various training outcomes. The relationship between training duration and the overall effect size was positive and significant, $b = .002$, $p = .002$ (see Table 3). Further, the *b* coefficients representing the relationships between hours of diversity training and specific training outcomes were all positive and statistically significant (cognitive learning, $b = .003$, $p = .000$; attitudinal/affective learning,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 4
Diversity Training Effects on Outcomes by Setting

Variable	Between-class effect *(Qb)*	*k*	*N*	*g* (*SE*)	*T*	95% CI for *g*
Cognitive learning						
Organizational		33	3,264	.52 (.08)	.38	[.33, .66]
Educational		71	5,650	.60 (.06)	.43	[.49, .71]
	.76					
Attitudes/Affective learning						
Organizational		55	5,206	.38 (.05)	.39	[.29, .47]
Educational		174	21,242	.28 (.03)	.24	[.23, .33]
	3.52†					
Behavioral learning						
Organizational		28	2,163	.51 (.07)	.26	[.38, .64]
Educational		65	5,962	.46 (.04)	.27	[.38, .55]
	.45					
Reactions						
Organizational		4	546	.28 (.19)	.21	[.01, .55]
Educational		8	890	.80 (.14)	.37	[.50, 1.09]
	6.43*					

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; *k* = the number of studies providing information included in the analysis; *N* = sum of the sample sizes of studies providing information included in the analysis; *g* (*SE*) = Hedges's *g* with standard error; *T* = indicator of heterogeneity; CI = confidence interval.
[a] For studies with multiple outcomes, the mean across outcomes is used in the overall effect analysis, whereas all of the selected outcomes, assuming independence, are used in the four evaluation criteria analysis.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

b = .002, *p* = .000; behavioral learning, *b* = .003, *p* = .016; reactions, *b* = .021, *p* = .000). The advantage of longer training interventions does seem to transfer to more positive reactions, and better diversity knowledge, attitudes, and skills, as reflected in the positive relationship between the hours of training and these outcome measures.

Types: Awareness and behavior based. We first analyzed differences in overall effect size as a function of diversity training types (see Table 2). As expected, we found that studies describing awareness-based training reported the lowest overall effect size (*g* = .31) compared to other types of diversity training (*g* = .46). The difference in effect sizes across types of diversity training was statistically significant, $Q_B(2) = 10.08$, *p* = .02. We further analyzed these data broken down by the type of outcome measured (see Table 8). Consistent with our theorizing, awareness training had smaller effect sizes for attitudinal/affective and behavioral learning (*g* = .22; *g* = .35, respectively) compared to other types of diversity training. The differences in effect sizes were statistically significant for attitudinal/affective learning, $Q_B(2) = 15.16$, *p* = .00 and behavioral learning, $Q_B(2) = 6.92$, *p* = .05, yet not significant for other outcomes.

Table 5
Diversity Training Effects on Outcomes by Approach

Variable	Between-class effect *(Qb)*	*k*	*N*	*g* (*SE*)	*T*	95% CI for *g*
Cognitive learning						
Integrated		9	699	.69 (.17)	.43	[.37, 1.01]
Standalone		95	8,215	.57 (.05)	.41	[.46, .65]
	.53					
Attitudes/Affective learning						
Integrated		32	3,137	.47 (.07)	.34	[.34, .63]
Standalone		194	22,727	.27 (.02)	.24	[.23, .33]
	7.15*					
Behavioral learning						
Integrated		13	1,442	.86 (.19)	.63	[.48, 1.23]
Standalone		80	6,683	.42 (.03)	.17	[.36, .48]
	5.11*					

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; *k* = the number of studies providing information included in the analysis; *N* = sum of the sample sizes of studies providing information included in the analysis; *g* (*SE*) = Hedges' g with standard error; *T* = indicator of heterogeneity; CI = confidence interval. Analyses for reaction measure were not possible because less than 2 data points were available for this measure.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 6
Diversity Training Effects on Outcomes by Attendance Requirements

Variable	Between-class effect *(Qb)*	*k*	*N*	*g* (*SE*)	*T*	95% CI for *g*
Cognitive learning						
Mandatory		15	1,141	.63 (.12)	.42	[.40, .87]
Voluntary		88	7,742	.56 (.05)	.40	[.46, .66]
	3.73					
Attitudes/Affective learning						
Mandatory		42	7,163	.36 (.05)	.27	[.25, .46]
Voluntary		174	17,952	.28 (.03)	.27	[.23, .34]
	1.66					
Behavioral learning						
Mandatory		21	2,349	.63 (.12)	.48	[.40, .86]
Voluntary		68	5,621	.42 (.03)	.17	[.35, .48]
	7.71*					
Reactions						
Mandatory		2	152	.37 (.30)	.32	[–.21, .96]
Voluntary		9	1,098	.71 (.14)	.38	[.43, .98]
	7.49*					

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; *k* = the number of studies providing information included in the analysis; *N* = sum of the sample sizes of studies providing information included in the analysis; *g* (*SE*) = Hedges' *g* with standard error; *T* = indicator of heterogeneity; CI = confidence interval.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Instruction: Many versus one method. We found no significant differences in overall effect sizes across types of instruction, $Q_B(1) = 2.85$, $p = .24$ (see Table 2). When analyzed by the type of outcome measured (see Table 9), we found that diversity training based on many instructional methods had larger effect sizes for reactions ($g = .73$) than studies that described training based on one instructional method ($g = .59$), $Q_B(1) = 6.91$, $p = .03$). The effects were in a similar direction for behavioral learning, yet did not reach statistical significance, $Q_B(1) = 2.46$, $p = .08$. No statistically significant differences in effect sizes across types of instruction were found for cognitive, $Q_B(1) = 1.07$, $p = .25$, or attitudinal/affective learning, $Q_B(1) = .28$, $p = .60$.

Trainee characteristics (Hypothesis 6). A metaregression analysis revealed that the average age of participants in the sample did not moderate the overall effect size, $b = .001$, $p = .70$ (see Table 3). No relationships were observed for cognitive ($b = .012$, $p = .25$), attitudinal/affective ($b = -.003$, $p = .51$), behavioral learning ($b = .003$, $p = .66$) or reactions ($b = -.009$, $p = .41$) when broken down by a specific outcome.

Table 7
Diversity Training Effects on Outcomes by Training Focus

Variable	Between-class effect *(Qb)*	*k*	*N*	*g* (*SE*)	*T*	95% CI for *g*
Cognitive learning						
Inclusive		10	1,158	.47 (.15)	.44	[.16, .76]
One group specific		81	6,909	.60 (.06)	.41	[.49, .71]
Multiple group specific		11	847	.53 (.13)	.36	[.27, .78]
	.62					
Attitudes/Affective learning						
Inclusive		34	5,139	.40 (.06)	.30	[.28, .52]
One group specific		167	18,084	.27 (.03)	.27	[.22, .33]
Multiple group specific		24	3,225	.30 (.07)	.31	[.16, .45]
	3.44					
Behavioral learning						
Inclusive		16	1,780	.53 (.08)	.25	[.37, .69]
One group specific		58	4,051	.51 (.04)	.25	[.43, .61]
Multiple group specific		11	1,147	.36 (.11)	.30	[.15, .58]
	1.78					

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; *k* = the number of studies providing information included in the analysis; *N* = sum of the sample sizes of studies providing information included in the analysis; *g* (*SE*) = Hedges' *g* with standard error; *T* = indicator of heterogeneity; CI = confidence interval. Analyses for reaction measure were not possible because less than 2 data points were available for this measure.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 8
Diversity Training Effects on Outcomes by Training Types

Variable	Between-class effect *(Qb)*	*k*	*N*	*g* (*SE*)	*T*	95% CI for *g*
Cognitive learning						
Awareness based		43	4,592	.60 (.07)	.42	[.46, .74]
Behavior based		—	—	—	—	—
Both		60	4,196	.57 (.06)	.41	[.44, .69]
	.12					
Attitudes/Affective learning						
Awareness based		116	16,547	.22 (.03)	.25	[.16, .27]
Behavior based		8	492	.41 (.10)	.10	[.22, .60]
Both		101	8,845	.40 (.03)	.28	[.32, .46]
	15.16***					
Behavioral learning						
Awareness based		29	2,859	.35 (.05)	.19	[.25, .45]
Behavior based		7	749	.53 (.12)	.26	[.29, .77]
Both		53	4,232	.54 (.06)	.36	[.42, .66]
	6.92*					

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; k = the number of studies providing information included in the analysis; N = sum of the sample sizes of studies providing information included in the analysis; *g* (*SE*) = Hedges' *g* with standard error; T = indicator of heterogeneity; CI = confidence interval. Analyses for reaction measure were not possible because less than 2 data points were available for this measure.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Similar analyses were conducted to examine the influence of gender composition in training groups (see Table 3). The relationship between the proportion of female participants in a sample and the overall effect size was positive but not statistically significant, $b = .087$, $p = .43$. When broken down by specific outcomes, this effect became statistically significant for reactions, $b = 1.86$, $p = .00$, indicating that larger effect sizes were associated with more women in training groups. No relationships were observed for cognitive ($b = .133$, $p = .66$), attitudinal/affective ($b = -.029$, $p = .82$), or behavioral learning ($b = -.079$, $p = .69$).

Finally, we examined whether the proportion of white participants in a sample predicted effect sizes (see Table 3). A metaregression analysis revealed that race of participants in the sample did not moderate the overall effect size, $b = .119$, $p = .28$. No relationship was observed for cognitive ($b = .125$, $p = .67$), attitudinal/affective learning ($b = .178$, $p = .17$), behavioral learning ($b = -.104$, $p = .59$) or reactions ($b = .960$, $p = .19$) when broken down by specific outcomes.

Study rigor (Hypothesis 7). We first analyzed differences in overall effect sizes across different study methodologies (*experi-*

Table 9
Diversity Training Effects on Outcomes by Instruction

Variable	Between-class effect *(Qb)*	*k*	*N*	*g* (*SE*)	*T*	95% CI for *g*
Cognitive learning						
One		23	2,286	.67 (.12)	.53	[.45, .90]
Many		79	6,628	.54 (.05)	.40	[.44, .65]
	1.07					
Attitudes/Affective learning						
One		71	9,724	.27 (.03)	.18	[.21, .34]
Many		151	16,724	.30 (.03)	.30	[.24, .36]
	.28					
Behavioral learning						
One		23	2,597	.39 (.06)	.19	[.29, .50]
Many		64	5,053	.51 (.05)	.30	[.42, .60]
	2.46†					
Reactions						
One		6	527	.59 (.08)	.11	[.43, .76]
Many		5	499	.73 (.25)	.54	[.23, 1.22]
	6.91*					

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; k = the number of studies providing information included in the analysis; N = sum of the sample sizes of studies providing information included in the analysis; *g* (*SE*) = Hedges' *g* with standard error; T = indicator of heterogeneity; CI = confidence interval.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

mental, *quasiexperimental*, and *observational* designs) and found no significant differences, $Q_B(2) = 3.05$, $p = .27$ (see Table 2). Further, effect sizes of true experiments with random assignment of participants to conditions were not significantly different from other designs when analyzing the effects for separate outcome variables, cognitive, $Q_B(2) = .49$, $p = .78$; attitudinal/affective, $Q_B(2) = 4.04$, $p = .13$; behavioral learning, $Q_B(2) = .78$, $p = .68$; reactions, $Q_B(2) = .67$, $p = .72$ (see Table 10).

Discussion

Given the extensive nature of this meta-analysis, we now try to answer some overall questions, among them: what did we learn about the body of diversity training research? What controversies within the field are we able to resolve? Do the diversity training characteristics widely accepted as effective matter as much as we think? Or are our ideas about what factors make diversity training effective misplaced? To do this, we first consider our findings on diversity training and learning outcomes, then we look at context, design and participants' role in shaping diversity training effects, before concluding with remarks on implications for educational policy and research. Our goal for this section, inspired in part by an anonymous reviewer, is to be as straightforward as possible to present a review for a broad audience, beyond those that research diversity issues, with information they might not otherwise encounter.

How Does Diversity Training Affect Learning Outcomes?

Overall, we found that the effects of diversity training vary as a function of the outcome used to operationalize its effectiveness. Consistent with our predictions, reactions to the training itself feature the strongest overall positive effects (see Table 11). This finding may be partly interpreted as due to demand characteristics—for example, "I might have liked the program because the trainer had a really great sense of humor, but I didn't change my diversity-related attitudes and behaviors one bit." Yet it is also possible for an inspiring or especially effective diversity trainer to bring about sustained, positive emotional responses in participants—for example, "That diversity instructor changed my life because it helped me in how I interact with people much older than me." Turning to other outcomes, and consistent with Kalinoski and colleagues' (2013) study, diversity training programs seem less effective in changing attitudes. This finding is in line with the notion that attitudes assessed in this research are generally strong (Eagly & Chaiken, 1993), particularly emotion-laden, tightly tied to trainees' self-identity, operate in highly connected networks, and are generally resistant to change (Dalege et al., 2016; Eagly & Chaiken, 2007; Kulik & Roberson, 2008a).

Comparing the immediate versus long-term effects of diversity training, we found that diversity training effects on reactions and attitudinal/affective learning decayed over time. This conclusion seems inconsistent with the few prior studies that have reported evidence of successful transfer of diversity training in the work setting (e.g., Hanover & Cellar, 1998; Majumdar, Browne, Roberts, & Carpio, 2004; Roberson, Kulik, & Pepper, 2009). We show that reactions and attitudes might be malleable for participants posttraining. For example, someone who is prejudiced against African Americans before taking diversity training may experience a positive shift in attitudes and become less prejudiced. Yet, their attitudes may shift back closer to what they were pretraining in response to media accounts of riots and unrest such as that occur-

Table 10
Diversity Training Effects on Outcomes by Study Rigor

Variable	Between-class effect *(Qb)*	*k*	*N*	*g* (*SE*)	*T*	95% CI for *g*
Cognitive learning						
Experimental		15	1,358	.71 (.22)	.82	.28/1.14
Quasi-experimental		47	3,888	.57 (.08)	.43	.41/.70
Observational		42	3,668	.55 (.06)	.34	.43/.67
	.49					
Attitudes/Affective learning						
Experimental		58	5,293	.26 (.05)	.28	.15/.36
Quasi-experimental		84	11,603	.37 (.04)	.32	.29/.46
Observational		85	9,552	.27 (.03)	.27	.21/.34
	4.04					
Behavioral learning						
Experimental		20	2,850	.54 (.10)	.33	.36/.74
Quasi-experimental		39	2,645	.44 (.08)	.39	.29/.59
Observational		34	2,630	.49 (.05)	.39	.40/.58
	.78					
Reactions						
Experimental		3	342	.48 (.12)	.00	.26/.71
Quasi-experimental		5	538	.63 (.31)	.66	.02/1.25
Observational		4	556	.66 (.21)	.39	.26/1.07
	.67					

Note. Between-class effect *(Qb)* = indicator of homogeneity that indicates whether the effect sizes differ among the levels of the categorical moderator; *k* = the number of studies providing information included in the analysis; *N* = sum of the sample sizes of studies providing information included in the analysis; *g* (*SE*) = Hedges' *g* with standard error; *T* = indicator of heterogeneity; CI = confidence interval.
$^{\dagger} p < .10$. $^{*} p < .05$. $^{**} p < .01$. $^{***} p < .001$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 11
A Summary of Results

Hypothesized relationships	Empirical support	Results
H1: Diversity training will have stronger effects on participants' reactions relative to cognitive, behavioral, and attitudinal learning.	Supported	Largest effect on reactions (g = .61), followed by cognitive (g = .57), behavioral (g = .48), and attitudinal/affective learning (g = .30)
H2: The effects of cognitive learning will persist, whereas the effects of attitudinal and behavioral learning will subside over time.	Supported	Attitudinal/affective learning and reactions to training decay, yet cognitive knowledge is maintained over time after diversity training
H3: Diversity training will have stronger effects on all learning outcomes when the training context provides more motivation to learn (e.g., educational settings, integrated, and mandatory) than when it does not (e.g., organizational settings, standalone, and voluntary).	Partially supported	Reactions to training were higher in *educational settings* Studies that utilized an *integrated approach* reported higher overall effect sizes (g = .57) than those that utilized a standalone approach (g = .36) *Mandatory* diversity training seemed more effective for behavioral learning, yet *voluntary* training was perceived more favorably by training participants
H4: Diversity training will have stronger effects on all learning outcomes when the design features provide more opportunities for cooperative contact (e.g., inclusive and longer) than when they do not (e.g., group specific and shorter).	Partially supported	No significant differences in effect sizes across *inclusive*, or *group-specific* diversity training *Longer* training interventions transfer to more positive reactions, and better diversity knowledge, attitudes, and skills
H5: Diversity training will have stronger effects on all learning outcomes when design features maximize learning opportunities (e.g., combined and multi-instructional) than when they do not (e.g., awareness or behaviour based and one instructional method).	Partially supported	*Awareness* diversity training had smaller overall effect size (g = .31) than other types of diversity training (g = .46) *Multi-instructional* training had larger effect sizes for reactions (g = .73) than training based on *one instructional* method (g = .59)
H6: Diversity training will have stronger effects on all learning outcomes when training groups consist of more women, minorities, and younger participants.	Partially supported	The average age, gender, or race of participants in the sample did not moderate the overall effect size Larger effect sizes in reactions were associated with more women in training groups
H7: Study rigor will be associated with higher effect sizes.	Not supported	There were no significant differences in overall effect sizes across different study methodologies

Note. H = Hypothesis.

ring in Ferguson or similar events, especially if the nature of such reports casts minorities in a negative light. So, as the first study that considered both short- and long-term effects of training, we find no compelling evidence that long-term effects of diversity training are sustainable in relation to attitudinal/affective outcomes.

In contrast, training effects on cognitive learning remained stable or in some cases even increased in the long-term (see Table 11). While we cannot fully explain the relative "stickiness" of cognitive learning, it may be that after training, cues in the workplace or elsewhere could reinforce cognitive responses that trainees learned. Perhaps people are reminded of scenarios or situations they have learned while in training, which then is more readily maintained and even strengthened over time. These cues could also come from mass media or other sources outside the immediate workplace or school. The influence of such sources on individual cognitions, sense-making, and even management decisions has been established by research based on management fashion theory (Abrahamson & Fairchild, 1999; Spell & Blum, 2005) as an important shaper of cognitions both in and out of the workplace.

What Role Do Context, Design, and Participants Play in Diversity Training?

Turning to specific features of diversity training itself, we found that some factors widely believed to determine effectiveness of diversity training (e.g., use of multiple training methods) did not lead to effectiveness. Yet, several features consistently associated with effective training (e.g., integrated training) are absent in the majority of programs. Overall, the effectiveness of diversity training varied as a function of diversity training context, design, and to a lesser degree the characteristics of trainees. These factors proved to be critical in shaping and moderating the main effects of diversity training on learning outcomes. Ultimately, our study shows that while many diversity training programs do not follow what we identify as best practices, the good news is that some programs do and guidelines for successful training are emerging from past research.

Training Context

Because we attempted to take a comprehensive approach to examining contextual factors behind diversity training effects, we are able to make some conclusions about their collective influences that can help resolve some outstanding questions in the diversity literature. In general, we find that the significance of a place or setting (i.e., *organizational* vs. *educational setting*) may have been overstated; the more important issue is the relationship between diversity training efforts and other initiatives that complement the training. Taking a closer look at setting, the difference in overall effect size as a function of diversity training setting approaches, but does not reach statistical significance. This is an important finding, because past reviews tend to focus on one or the other rather than a combination of settings (Arai, Wanca-Thibault,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

& Shockley-Zalabak, 2001). However, consistent with our predictions, reactions to training were higher in educational settings, which we believe is because diversity training is more easily seen as part of the overall curriculum, and not something that "takes time away from work."

Our results further showed that diversity training has been most useful when training was *integrated* or embedded (as opposed to standalone). Indeed, as we predicted integrated efforts may signal managerial commitment to diversity above and beyond that of a single class or seminar, substantially increasing the motivation of participants to learn. Integrated training also means components are more likely to complement or support one another. For example, a social networking group of minority professionals, supported by the organization, is a follow-up outcome of a diversity course and also serves as a mentoring source. As our study is the first to include this contextual factor in quantitative analysis, the strong effects we found reveal the criticality of offering diversity programs as part of a well-thought out package or portfolio of diversity-related efforts (see Table 11). By showing us just how important it is relative to other factors, it resolves prior questions in diversity training research about "what matters most?" in terms of contextual choices.

Turning to attendance requirements, while an overall effect was not significant, when looking at separate outcomes, *mandatory* diversity training seemed more effective for behavioral learning, yet *voluntary* training was perceived more favorably by training participants. This latter effect may be due to selection aspects—people who willingly take training already have an interest in the issue (or they would not have volunteered) and are more likely to enjoy diversity training. Seen another way, they would be less likely to bring with them negative ideologically based biases ("this will all be politically correct propaganda") than people forced to take training. Our findings also reflect an interesting controversy in the literature with respect to whether mandatory diversity training should be more effective than voluntary training or vice versa. While participants like to have a choice (Dobbin & Kalev, 2007; Kaplan, 2006), Pettigrew and Tropp's (2006) research based on the contact perspective suggests that the voluntary approach does not lead to the strongest effects. One reason for this could be that under the voluntary scenario, people participating in training already want to be there and are not necessarily the ones who would benefit most from changes in cognitive, attitudinal, or behavioral outcomes (Ellis & Sonnenfeld, 1994).

Training Design

Unlike what we predicted based on psychological theory on diversity, we did not find any strong effects with respect to the *focus* of diversity training (i.e., inclusive, one group, or multiple groups). Whether focusing on one or more groups (e.g., the experiences of African Americans, then women, then gay lesbian bisexual and transgender, then persons with disabilities, etc.) or discussing more generic issues such as ingroup versus outgroup dynamics that may be generalizable to all types of demographic differences does not really matter or explains any additional variation in the effects of diversity training. What matters most, however, is the length of diversity training. We found a strong and significant relationship between the length of diversity training and effect sizes suggesting that diversity training programs that are *longer* tend to be more effective (see Table 11). Psychological theory on diversity and quantitative aspects of contact hypothesis predict longer programs provide more opportunities for contact. This is also consistent with training literature suggesting that more practice leads to greater skill development (e.g., Ericsson, Krampe, & Tesch-Römer, 1993).

Most effective types of diversity training programs were primarily designed to increase *both* diversity awareness and skills. The overall effect across different types of diversity training (awareness only, behavior only, and both awareness and behavior) was strong and significant, as was the effect for attitudes/affective and behavioral learning when analyzed for separate outcome measures. Although some authors questioned the inclusion of a behavioral component at all, especially in lesbian, gay, bisexual, and transgender diversity training where there is the potential for backlash from some employees who believe that such sexual orientations are inherently wrong (Kaplan, 2006), it seems that awareness and behavior-based diversity training is more effective when done together than separately. While it should be noted that our results showed training coded as behavior only was also effective, in retrospect it is hard to imagine training having no awareness element. Theoretically, these results make sense because making people aware of an issue or need for changing behavior would increase the likelihood of behavioral changes in response. Finally, unlike what we expected based on prior training research, we did not find any strong effects with respect to training instruction (i.e., one or many methods), except for the reaction measure. It seems that trainees tend to respond more favorably to programs that employ *many* instructional methods (e.g., lectures, exercises, group activities and discussions, etc. all together).

Trainees

We did not find any significant effects related to demographic characteristics of trainees except that a greater proportion of women in training groups seems to be related to more favorable responses to the training itself (see Table 11). This effect makes sense in that women tend to be more receptive of diversity training (e.g., Butler, Ryan, & Juarez, 2012). That race is not related to effectiveness may also be explained by past research (Roberson et al., 2001) that found the race–effectiveness relationship was dependent on context, prior experience of trainees with diversity training (we do not know the past experience of each person taking training), and training design. Given the breadth of our sample, the effects may simply cancel each other out with respect to the effects of race. Likewise, no relationship was observed for the age composition of training groups, even though age has been implicated as a primary shaper in numerous social trends, from voting patterns to attitudes about entitlements (Pew Research Center, 2012). The reason for this noneffect may be that contextual factors, as for race, cancel out any relationships.

Study Rigor

Turning to the role of *study rigor* in the studies we analyzed, we found that contrary to our expectations and a substantial body of research in social psychology, we do not see any significant differences between experimental and nonexperimental work. This unexpected result further supports a general theme we uncovered

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

in this analysis. Despite differences in methodological traditions that different disciplines studying diversity training have followed, and the presumed impacts of those approaches, we find that it is the content of the training that matters most (e.g., whether it is embedded, length of the training) and not factors like empirical approach of the study and its setting. In other words, what is included in the delivery of the training, and how it is supported by the organization, and how much that content motivates participants seems more important than where the training is conducted or how data were collected.

Educational and Policy Implications

These findings have important implications for educational institutions, particularly in light of APA guidelines, learning goals for undergraduate psychology majors and others. Such guidelines highlight the importance of developing socially responsible behaviors for professional and personal settings in a landscape that involves increasing diversity. In fact, scholars have made the assertion that diversity training and learning about diverse groups of people should be a learning outcome for all students (Dunn et al., 2010). As our meta-analytical review shows, the good news is that cognitive learning, presumably a key focus of the educational arena, was an outcome found to be more likely to persist or even increase over time. Our results also demonstrate that larger effect sizes were associated with integrated diversity training programs—and this is also consistent with the APA guidelines that encourage an intense, immersive experience and integrated approaches to diversity components in the undergraduate curriculum.

To that end, our comprehensive meta-analytical integration of research on the effects of diversity training programs demonstrates that certain contextual (e.g., training approach) and design factors (e.g., training type) of diversity training procedures had larger effect sizes overall and could lead to distinct diversity-related outcomes for participants. Cognitive learning, as we mentioned above, was a more persistent training outcome than perception, attitude or behavioral change associated with the training. Understanding how diversity training programs influence different learning outcomes represents an important consideration in any program review or development. Beyond the importance of training approach and type, more structural and mechanical considerations of the training have policy implications—essentially, volume matters. Length of the training was one of the characteristics related to larger effect sizes and this can inform policy decision about how much time should be devoted to training, especially in educational settings where semester or year-long efforts are typically undertaken.

Our results can also inform policy connected to diversity training and diversity-related issues. Cases such as the gay suicide case at Rutgers, the Ferguson shooting and related protests, and religious based conflict in France have clear implications for policy on both the level of individual organizations but also society as a whole. As diversity training focuses on serious, unresolved social problems, such as race and gender relations (Fowler, 2006) it can be seen as a response to these events, as well as to local incidents in schools and workplaces. Diversity training can work as an "on the ground" approach by providing tools, information, and knowledge to help employees and students not only understand these societal issues but also apply them in day to day interactions with those of another race, religion, or ethnic group.

Implications for Research

This meta-analysis provides a comprehensive review of research on diversity training. Nevertheless, it does not address every pertinent question related to this topic. Some questions may be so fine-grained that the sample sizes of relevant studies would be restricted to an unacceptably small size. For example, the instances where studies measured reactions to training were limited, which was somewhat surprising given research reported reactions as one of more the popular outcome measures of training (e.g., Curtis et al., 2007). However, we were able to include this important outcome measure in some of our analyses. In addition, while our results suggest that demographic and other characteristics of the instructor may go a long way toward shaping diversity training outcomes, fine-grained analyses were not possible because many studies did not report such information. Given the fact that the number of diversity training studies has exploded in recent years and thus may allow for such analysis soon, one avenue of possible research would be to more closely explore this issue.

Even with the above limitations, the cumulative empirical evidence we report has provided many insights into our understanding of diversity training programs. A key finding from our analysis is that integrated training worked well along with training that focused on both skill-building and awareness. From these conclusions a question arises: what exactly needs to be integrated? Our findings support a systems approach to diversity management (Curtis et al., 2007) where the configuration of programs and methods would be a critical part of diversity training effectiveness. Aside from this general conclusion, we still do not know the nature of successful configurations of initiatives surrounding a diversity training program, or even how many programs should be in place (e.g., supervisor support training, initiatives to explicitly convey commitment of the organization to diversity). Future research needs to determine the critical mass and nature of supporting programs for diversity training to be effective.

In contrast, the lack of findings with respect to training focus could also inspire some further investigation. For example, focusing on multiple groups versus discussing more generic issues such as ingroup versus outgroup dynamics could be both equally important and provide a lens to reflect how the world, and many (most) workplaces operate. Workplaces are increasingly represented by numerous demographic groups and often times these groups work and socialize separately. In fact, faultline theory (Lau & Murnighan, 1998; Thatcher & Patel, 2012) is based on divisions within groups of people that are aligned based on multiple demographic characteristics. A growing amount of empirical research has established the prevalence of faultlines in numerous educational and work contexts (e.g., Bezrukova, Jehn, Zanutto, & Thatcher, 2009; Jehn & Bezrukova, 2010). Yet, faultlines are embedded in systems, groups, and organizations that can provide a common ingroup identity to produce positive feelings toward ingroups (Tajfel & Turner, 1986). Indeed the effectiveness of training that takes such systems view, with a perspective that recognizes the complexities of the world, has been widely accepted by organizational researchers (Garvin, Cullen, & Datar, 2010; Waddock & Lozano, 2013). How we can model this dynamic

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

between distinct subgroup and overall group identifications in diversity training programs becomes an interesting research question.

Research on diversity training and education has been also criticized due to a number of methodological limitations including the "demand characteristic" problem (e.g., people are motivated to look good so they give the "right" answers). Because this is an important concern that readers (including researchers who conduct research on diversity training) should be aware of, we have additionally coded whether studies used explicit (e.g., Multicultural Awareness, Knowledge, and Skills Survey, D'Andrea et al., 1991; the Modern Racism Scale, McConahay, Hardee, & Batts, 1981) or implicit (e.g., latency-based measurements such as the Implicit Association Test or objective measures such as peer or supervisor appraisal ratings) measures of attitudes and behaviors. Our supplemental analysis shows that while the effect size decreased from .32 to .21 (95% CI [0.08, 0.35]), the value was still significant ($p < .01$), suggesting that the reported mean effect is not an artifact of outcome measurement. So, while we acknowledge the potential for demand characteristics to alter interpretations, we found no evidence that this changes the interpretations of our results. However, we do recommend using more implicit measures in future research when assessing the outcomes of diversity training.

In fact, methodological changes that need to be implemented in future studies to address the "demand characteristic" concern include the development of assessment instruments based on implicit measures of attitudes and behaviors (e.g., the Implicit Association Test, Greenwald et al., 1998; Haines & Sumner, 2006; The Instructor Cultural Competence Questionnaire, Roberson, Kulik, & Pepper, 2002). Although these instruments are not entirely free of the demand problem, they may allow generation of yet-untapped assessments of important preferences, attitudes, stereotypes, and behaviors which would advance the evaluation of diversity training. An even better method is to devise dependent measures that the participants perceive as unrelated to the experiment. For example, more objective measures (e.g., application of peers' evaluations, observations, and behavioral/physiological measures in addition to traditional direct self-reports) should be considered. Result-oriented measures such as statistics on discrimination lawsuits, Equal Employment Opportunity filings, grievances, turnover, bonuses, minority recruitment, and so forth in organizations undertaking diversity training may also be beneficial yet are rarely encountered.

In conclusion, we are happy to report that, contrary to charges made by our predecessors over the years, diversity training research is no longer atheoretical, irrelevant, or dull. This body of work has been very instrumental in offering theoretical guidance and insights into our understanding of the methods, process, and outcomes of diversity training (e.g., Paluck, 2006; Pendry et al., 2007; Wiethoff, 2004). Exciting advances in the areas of learning and motivation within the diversity training literature have been realized. We are now moving this work forward by integrating theoretical traditions from training literature with psychological theory on diversity. We thus, conclude on an optimistic note by reiterating that there are several exciting opportunities for future research in this area because (a) several theoretical perspectives have gained prominence in recent times to guide research, (b) the methodologies to adequately evaluate training effects have received attention in the literature, and (c) the studies reviewed above lay a groundwork for future research that has the opportunity to impact organizations, and society, in a demonstrable and positive way.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abernethy, A. D. (2005). Increasing the cultural proficiency of clinical managers. *Multicultural Counseling and Development, 33,* 81–93. http://dx.doi.org/10.1002/j.2161-1912.2005.tb00007.x

Abrahamson, E., & Fairchild, G. (1999). Management fashions: Lifecycles, triggers, and collective learning processes. *Administrative Science Quarterly, 44,* 708–740. http://dx.doi.org/10.2307/2667053

Aguinis, H., & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations, and society. *Annual Review of Psychology, 60,* 451–474. http://dx.doi.org/10.1146/annurev.psych.60.110707.163505

Alderfer, C. P. (1992). Changing race relations embedded in organizations: Report on a long-term project with the XYZ corporation. In S. E. Jackson (Ed.), *Diversity in the workplace: Human resource initiatives* (pp. 138–166). New York, NY: Guilford Press.

*Aldridge, S. W. (2013). *The impact of generational awareness training on teacher collaboration* (Unpublished dissertation). Trevecca Nazarene University, Nashville, TN.

Allport, G. (1954). *The nature of prejudice*. Cambridge, MA: Addison Wesley.

*Alonso, E. (2005). *Can we teach multicultural competence? The effects of multicultural counseling training on the multicultural counseling competence of counselor trainees* (Unpublished dissertation). University of La Verne, La Verne, CA.

*Altshuler, L., Sussman, N. M., & Kachur, E. (2003). Assessing changes in intercultural sensitivity among physician trainees using the Intercultural Development Inventory. *International Journal of Intercultural Relations, 27,* 387–401. http://dx.doi.org/10.1016/s0147-1767(03)00029-4

*Amatea, E., Cholewa, B., & Mixon, K. (2012). Influencing preservice teachers' attitudes about working with low-income and/or ethnic minority families. *Urban Education, 47,* 801–834. http://dx.doi.org/10.1177/0042085912436846

American Psychological Association (APA). (2013). *APA guidelines for the undergraduate psychology major: Version 2.0*. Retrieved from http://www.apa.org/ed/precollege/undergrad/index.aspx

Anand, R., & Winters, M. (2008). A retrospective view of corporate diversity training from 1964 to the present. *Academy of Management Learning & Education, 7,* 356–372. http://dx.doi.org/10.5465/AMLE.2008.34251673

*Anderson, C. L. (1981). The effect of a workshop on attitudes of female nursing students toward male homosexuality. *Journal of Homosexuality, 7,* 57–69. http://dx.doi.org/10.1300/J082v07n01_06

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., . . . Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin, 136,* 151–173. http://dx.doi.org/10.1037/a0018251

Anderson, J. W. (1982). *A study to determine the effects of facilitating conditions and action-oriented techniques in cross-cultural counseling* (Unpublished dissertation). University of Illinois at Urbana–Champaign, Champaign, IL.

Arai, M., Wanca-Thibault, M., & Shockley-Zalabak, P. (2001). Communication theory and training approaches for multiculturally diverse organizations: Have academics and practitioners missed the connection? *Public Personnel Management, 30,* 445–455. http://dx.doi.org/10.1177/009102600103000402

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

*Arizaga, M. P. (1999). *A Multicultural Relationship Enhancement Approach to Remediation and Prevention of Prejudice* (Unpublished dissertation). New Mexico State University, Las Cruces, NM.

*Armour, M. P., Bain, B., & Rubio, R. (2004). Special section: Field education in social work. An evaluation study of diversity training for field instructors: A collaborative approach to enhancing cultural competence. *Journal of Social Work Education, 40,* 27–38.

Arthur, W., Bennett, W., Jr., & Huffcutt, A. I. (2001). *Conducting meta-analysis using SAS*. Mahwah, NJ: Erlbaum.

Arthur, W., Jr., Bennett, W., Jr., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88,* 234–245. http://dx.doi.org/10.1037/0021-9010.88.2.234

Avery, D. R., & Thomas, K. M. (2004). Blending content and contact: The role of diversity curriculum and campus heterogeneity in fostering diversity management competency. *Academy of Management Learning & Education, 3,* 380–396. http://dx.doi.org/10.5465/AMLE.2004.15112544

*Baba, Y., & Hebert, C. (2004). The effects of participation in a cultural awareness program on jail inmates. *Journal of Ethnic & Cultural Diversity in Social Work, 13,* 91–113. http://dx.doi.org/10.1300/J051v13n03_05

*Bailey, A., Barr, O., & Bunting, B. (2001). Police attitudes toward people with intellectual disability: An evaluation of awareness training. *Journal of Intellectual Disability Research, 45,* 344–350. http://dx.doi.org/10.1046/j.1365-2788.2001.00339.x

*Baker, S. A. (2009). *Recognizing bias related to diversity: An experimental research study* (Unpublished dissertation). Alliant International University, Fresno, CA.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and direction for future research. *Personnel Psychology, 41,* 63–105. http://dx.doi.org/10.1111/j.1744-6570.1988.tb00632.x

*Ballou, C. D. (1996). *A comparative assessment of the effectiveness of international/intercultural education in the general education program at American University* (Unpublished dissertation). American University, Washington, DC.

*Barker, S., & Hartel, C. E. J. (2004). Intercultural service encounters: An exploratory study of customer experiences. *Cross Cultural Management: An International Journal, 11,* 3–14. http://dx.doi.org/10.1108/13527600410797710

*Bauer, M., McAulife, L., Nay, R., & Chenco, C. (2013). Sexuality in older adults: Effect of an education intervention on attitudes and beliefs of residential aged care staff. *Educational Gerontology, 39,* 82–91. http://dx.doi.org/10.1080/03601277.2012.682953

Bendick, M., Egan, M. L., & Lofhjelm, S. (2001). Workplace diversity training: From anti-discriminatory compliance to organizational development. *Human Resource Planning, 24,* 10–25.

Bezrukova, K., Jehn, K., & Spell, C. (2012). Reviewing diversity training: Where we have been and where we should go. *Academy of Management Learning & Education, 11,* 207–227. http://dx.doi.org/10.5465/amle.2008.0090

Bezrukova, K., Jehn, K. A., Zanutto, E., & Thatcher, S. M. B. (2009). Do workgroup faultlines help or hurt? A moderated model of faultlines, team identification, and group performance. *Organization Science, 20,* 35–50. http://dx.doi.org/10.1287/orsc.1080.0379

*Bhawuk, D. P. S. (1998). The role of culture theory in cross-cultural training: A multimethod study of culture-specific, culture-general, and culture theory-based assimilators. *Journal of Cross-Cultural Psychology, 29,* 630–655. http://dx.doi.org/10.1177/0022022198295003

*Blakely, G. L., Blakely, E. H., & Moorman, R. H. (1998). The effects of training on perceptions of sexual harassment allegations. *Journal of Applied Social Psychology, 28,* 71–83. http://dx.doi.org/10.1111/j.1559-1816.1998.tb01654.x

*Bloch, C. (2012). *Cultural competency intervention program for healthcare workers* (Unpublished dissertation). Azusa Pacific University, Azusa, CA.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis* (Version 2) [Computer software]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, United Kingdom: Wiley. http://dx.doi.org/10.1002/9780470743386

*Boswell, S. S. (2012). "Old people are cranky": Helping professional trainees' knowledge, attitudes, aging anxiety, and interest in working with older adults. *Educational Gerontology, 38,* 465–472. http://dx.doi.org/10.1080/03601277.2011.559864

Braham, J. (1989). No, you don't manage everyone the same. *Industry Week, 238,* 28–35.

*Brathwaite, A. C., & Majumdar, B. (2006). Evaluation of a cultural competence educational programme. *Journal of Advanced Nursing, 53,* 470–479. http://dx.doi.org/10.1111/j.1365-2648.2006.03742.x

Brewer, M. B., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M. B. Brewer (Eds.), *Group in contact: The psychology of desegregation* (pp. 281–302). Orlando, FL: Academic Press. http://dx.doi.org/10.1016/B978-0-12-497780-8.50019-X

*Brown, E. L. (2004). What precipitates change in cultural diversity awareness during a multicultural course: The message or the method? *Journal of Teacher Education, 55,* 325–340. http://dx.doi.org/10.1177/0022487104266746

*Brown, S. P., Parham, T. A., & Yonker, R. (1996). Influence of a cross-cultural training course on racial identity attitudes of white women and men: Preliminary perspectives. *Journal of Counseling and Development, 74,* 510–516. http://dx.doi.org/10.1002/j.1556-6676.1996.tb01902.x

*Burris, R. M. (1992). *Measuring the learning outcomes of a continuing education seminar about the aging process on the knowledge level of registered nurses* (Unpublished dissertation). University of North Texas, Denton, TX.

*Bush, V., & Ingram, T. (2001). Building and assessing cultural diversity skills: Implications for sales training. *Industrial Marketing Management, 30,* 65–76. http://dx.doi.org/10.1016/S0019-8501(99)00089-9

*Butler, J., Ryan, C., & Juarez, C. (2012). *Effects of managerial support and rationale on diversity training effectiveness*. Paper presented at the Society for Industrial & Organizational Psychology, San Diego, CA.

*Byington, K., Fischer, J., Walker, L., & Freedman, E. (1997). Evaluating the effectiveness of a multicultural counseling ethics and assessment training. *Journal of Applied Rehabilitation Counseling, 28,* 15–19.

Byrne, D., & Wong, T. J. (1962). Racial prejudice, interpersonal attraction, and assumed dissimilarity of attitudes. *Journal of Abnormal and Social Psychology, 65,* 246–253. http://dx.doi.org/10.1037/h0047299

*Caffrey, R. A., Neander, W., Markle, D., & Stewart, B. (2005). Improving the cultural competence of nursing students: Results of integrating cultural content in the curriculum and an international immersion experience. *The Journal of Nursing Education, 44,* 234–240.

Callahan, J. S., Kiker, D. S., & Cross, T. (2003). Does method matter? A meta-analysis of the effects of training method on older learner training performance. *Journal of Management, 29,* 663–680. http://dx.doi.org/10.1016/S0149-2063(03)00029-1

*Cap, I. (1995). *A study of the usefulness and effectiveness of a self-instructional print module on multicultural behavior change in apprentices in Manitoba* (Unpublished dissertation). Florida State University, Tallahassee, FL.

*Carpenter, C. C. (1997). *Effects of interpersonal skills training on nurse aides' client care* (Unpublished dissertation). TX Woman's University, Denton, TX.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

*Carrell, L. J. (1997). Diversity in the communication curriculum: Impact on student empathy. *Communication Education, 46,* 234–244. http://dx.doi.org/10.1080/03634529709379098

*Carter, M. M., Lewis, E. L., Sbrocco, T., Tanenbaum, R., Oswald, J. C., Sykora, W., . . . Hill, L. D. (2006). Cultural competency training for third-year clerkship students: Effects of an interactive workshop on student attitudes. *Journal of the National Medical Association, 98,* 1772–1778.

Carver, C. S., & Scheir, M. F. (1990). Origins and functions of positive and negative affect: A control process view. *Psychological Review, 97,* 19–35. http://dx.doi.org/10.1037/0033-295X.97.1.19

*Cascio, W. F., & Bass, B. M. (1976). The effects of role playing in a program to modify attitudes toward black employees. *The Journal of Psychology, 92,* 261–266. http://dx.doi.org/10.1080/00223980.1976.9921365

*Case, K. A. (2007). Raising white privilege awareness and reducing racial prejudice: Assessing diversity course effectiveness. *Teaching of Psychology, 34,* 231–235. http://dx.doi.org/10.1080/00986280701700250

*Case, K. A., & Stewart, B. (2010). Changes in diversity course student prejudice and attitudes toward heterosexual privilege and gay marriage. *Teaching of Psychology, 37,* 172–177. http://dx.doi.org/10.1080/00986283.2010.488555

*Case, K. A., & Stewart, B. (2013). Intervention effectiveness in reducing prejudice against transsexuals. *Journal of LGBT Youth, 10,* 140–158. http://dx.doi.org/10.1080/19361653.2012.718549

*Castillo, L. G., Brossart, D. F., Reyes, C. J., Conoley, C. W., & Phoummarath, M. J. (2007). The influence of multicultural training on perceived multicultural counseling competencies and implicit racial prejudice. *Journal of Multicultural Counseling and Development, 35,* 243–254. http://dx.doi.org/10.1002/j.2161-1912.2007.tb00064.x

*Cates, J. T. (2006). *The effects of infusing multicultural training into counseling practicum on trainees' self-perceived multicultural counseling competence* (Unpublished dissertation). University of Nevada, Reno, NV.

*Celik, H., Abma, T. A., Klinge, I., & Widdershoven, G. A. M. (2012). Process evaluation of a diversity training program: The value of a mixed method strategy. *Evaluation and Program Planning, 35,* 54–65. http://dx.doi.org/10.1016/j.evalprogplan.2011.07.001

*Chang, M. J. (2002). The impact of an undergraduate diversity course requirement on students' racial views and attitudes. *The Journal of General Education, 51,* 1–42. http://dx.doi.org/10.1353/jge.2002.0002

*Chrobot-Mason, D. (2004). Managing racial differences: The role of majority managers' ethnic identity development on minority employee perceptions of support. *Group & Organization Management, 29,* 5–31. http://dx.doi.org/10.1177/1059601103252102

*Clark, M. A. (1998). *The effects of a cross-cultural mentoring program and diversity training on preservice teachers' attitudes and beliefs regarding diverse students* (Unpublished dissertation). University of Florida, Gainesville, FL.

Cleveland, J. N., Festa, R. M., & Montgomery, L. (1988). Applicant pool composition and job perceptions: Impact of decisions regarding an older applicant. *Journal of Vocational Behavior, 32,* 112–125. http://dx.doi.org/10.1016/0001-8791(88)90009-7

*Clevenger, K. (2011). *Effects of a formal service-learning program on baccalaureate nursing student's perception of their level of cultural competence* (Unpublished dissertation). Walden University, Minneapolis, MN.

*Clinton, L. T. (1983). *The effectiveness of two methods of teaching sociology in facilitating students' rational thinking about racial and ethnic groups* (Unpublished dissertation). East Texas State University, Commerce, TX.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

*Cole, E. R., Case, K. A., Rios, D., & Curtin, N. (2011). Understanding what students bring to the classroom: Moderators of the effects of diversity courses on student attitudes. *Cultural Diversity & Ethnic Minority Psychology, 17,* 397–405. http://dx.doi.org/10.1037/a0025433

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85,* 678–707. http://dx.doi.org/10.1037/0021-9010.85.5.678

*Colvin-Burque, A., Davis-Maye, D., & Zugazaga, C. (2007). Can cultural competence be taught? Evaluating the impact of the SOAP Model. *Journal of Social Work Education, 43,* 223–241. http://dx.doi.org/10.5175/JSWE.2007.200500528

*Combs, G. M., & Luthans, F. (2007). Diversity training: Analysis of the impact of self-efficacy. *Human Resource Development Quarterly, 18,* 91–120. http://dx.doi.org/10.1002/hrdq.1193

Cornelius, N., Gooch, L., & Todd, S. (2000). Managers leading diversity for business excellence. *Journal of General Management, 25,* 67–78.

*Cornett-DeVito, M., & McGlone, E. (2000). Multicultural communication training for law enforcement officers: A case study. *Criminal Justice Policy Review, 11,* 234–253.

Cox, T., Jr., & Blake, S. (1991). Managing cultural diversity: Implications for organizational competitiveness. *The Academy of Management Executive, 5,* 45–56. http://dx.doi.org/10.5465/AME.1991.4274465

Curtis, E. F., Dreachslin, J. L., & Sinioris, M. (2007). Diversity and cultural competence training in health care organizations: Hallmarks of success. *The Health Care Manager, 26,* 255–262. http://dx.doi.org/10.1097/01.HCM.0000285018.18773.ed

Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. J. (2016). Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychological Review, 123,* 2–22. http://dx.doi.org/10.1037/a0039802

*D'Andrea, M., Daniels, J., & Heck, R. (1991). Evaluating the impact of multicultural counseling training. *Journal of Counseling and Development, 70,* 143–150. http://dx.doi.org/10.1002/j.1556-6676.1991.tb01576.x

*Darnell, D. A., & Cook, S. L. (2009). Investigating the utility of the film War Zone in the prevention of street harassment. *Psychology of Women Quarterly, 33,* 266–274.

*Darst, E. H. (1988). *Effect of sex education on nurses' skills and clinical outcomes with disabled patients* (Unpublished dissertation). University of Missouri, Kansas City, KS.

*Davidson, J. R. (1991). *Evaluation of an education model for race/ethnic sensitive social work and critique of the white racial identity attitude scale* (Unpublished dissertation). University of Texas at Arlington, Arlington, TX.

*de Lemus, S., Navarro, L., Megias, J. L., Velasquez, M., & Ryan, E. (2013). *From sex to gender: A university intervention to reduce sexism in Argentina, Spain, and El Salvador*. Paper presented at the Society of Experimental Social Psychology, Berkeley, CA.

*Delgado, D. A., Ness, S., Ferguson, K., Engstrom, P. L., Gannon, T. M., & Gillett, C. (2013). Cultural competence training for clinical staff: Measuring the effect of a one-hour class on cultural competence. *Journal of Transcultural Nursing, 24,* 204–213. http://dx.doi.org/10.1177/1043659612472059

Del Portillo, C. T. (1981). *The effectiveness of minimal success on positive attitude change among Latino women* (Unpublished dissertation). University of San Francisco, San Francisco, CA.

*De Meuse, K. P., Hostager, T. J., & O'Neill, K. S. (2007). A longitudinal evaluation of senior managers' perceptions and attitudes of a workplace diversity training program. *Human Resource Planning, 30,* 38–46.

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

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

*Diaz-Lazaro, C., & Cohen, B. B. (2001). Cross-cultural contact in counseling training. *Journal of Multicultural Counseling and Development, 29,* 41–56. http://dx.doi.org/10.1002/j.2161-1912.2001.tb00502.x

*Dickson, G. L., Argus-Calvo, B., & Tafoya, N. G. (2010). Multicultural counselor training experiences: Training effects and perceptions of training among a sample of predominately Hispanic students. *Counselor Education and Supervision, 49,* 247–265. http://dx.doi.org/10.1002/j.1556-6978.2010.tb00101.x

Digh, P. (1998). The next challenge: Holding people accountable. *HR Magazine, 43,* 63–69.

Dobbin, F., & Kalev, A. (2007). *Kicking and screaming: Commitment and resistance in organizational diversity programs*. Paper presented at the American Psychological Association National Conference, San Francisco, CA.

*Dogra, N. (2001). The development and evaluation of a programme to teach cultural diversity to medical undergraduate students. *Medical Education, 35,* 232–241. http://dx.doi.org/10.1046/j.1365-2923.2001.00734.x

*Dorfman, L. T., Murty, S. A., Ingram, J. G., & Li, H. (2007). Evaluating the outcomes of gerontological curriculum enrichment: A multi-method approach. *Gerontology & Geriatrics Education, 27,* 1–21. http://dx.doi.org/10.1300/J021v27n04_01

*Drwecki, B. B. (2011). *An experimental investigation of the causes of and solutions to racial pain treatment bias* (Unpublished dissertation). Madison, WI: University of Wisconsin–Madison.

Dunn, D. S., Brewer, C. L., Cautin, R. L., Gurung, R. A., Keith, K. D., McGregor, L. N., . . . Halpern, D. F. (2010). The undergraduate psychology curriculum: Call for a core. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (pp. 47–61). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/12063-003

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455–463. http://dx.doi.org/10.1111/j.0006-341X.2000.00455.x

*Dyson, M. E. (2003). *The effectiveness of using a brief "in-group bias" training program: A multiple case study* (Unpublished dissertation). Temple University, Philadelphia, PA.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt, Brace, Jovanovich.

Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition, 25,* 582–602. http://dx.doi.org/10.1521/soco.2007.25.5.582

*Earley, P. C. (1987). Intercultural training for managers: A comparison of documentary and interpersonal methods. *Academy of Management Journal, 30,* 685–698. http://dx.doi.org/10.2307/256155

*Edwards, B. J. (1999). *The effect of multicultural training on the cross-cultural adaptability of college students* (Unpublished dissertation). Nova Southeastern University, Ft. Lauderdale, FL.

*Edwards, S. L. (1997). *Teaching strategies for multicultural competence* (Unpublished dissertation). University of South Carolina, Columbia, SC.

Ellis, K., & Sonnenfeld, J. (1994). Diverse approaches to managing diversity? *Human Resource Management, 33,* 79–109. http://dx.doi.org/10.1002/hrm.3930330106

*Ellison, H. Y. (2002). *The efficacy of the Ellison model as a retention initiative for first semester freshmen* (Unpublished dissertation). Florida International University, Miami, FL.

Ely, R. J. (2004). A field study of group diversity, participation in diversity education programs, and performance. *Journal of Organizational Behavior, 25,* 755–780. http://dx.doi.org/10.1002/job.268

Ensari, N. K., & Miller, N. (2006). The application of the personalization model in diversity management. *Group Processes & Intergroup Relations, 9,* 589–607. http://dx.doi.org/10.1177/1368430206067679

Epstein, A. (1994). Diversity training goes too far—FAA to pay workers who suffered after anti-harassment program. Retrieved from http://community.seattletimes.nwsource.com/archive/?date=19940923&slug=1932249

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406. http://dx.doi.org/10.1037/0033-295X.100.3.363

*Finken, L. L. (2002). The impact of a human sexuality course on anti-gay prejudice: The challenge of reaching male students. *Journal of Psychology & Human Sexuality, 14,* 37–46. http://dx.doi.org/10.1300/J056v14n01_03

*Fischer, R. (2011). Cross-cultural training effects on cultural essentialism beliefs and cultural intelligence. *International Journal of Intercultural Relations, 35,* 767–775. http://dx.doi.org/10.1016/j.ijintrel.2011.08.005

*Flamini, B. (2005). *The influence of multicultural training on the cross-cultural adaptability of university business students* (Unpublished dissertation). Nova Southeastern University, Ft. Lauderdale, FL.

*Fluck, U. (2003). *Development of an online supplement to reduce ethnocentrism* (Unpublished dissertation). University of Montana, Missoula, MT.

Foldy, E. G., & Creed, D. W. E. (1999). Action learning, fragmentation, and the interaction of single-, double-, and triple-loop change. *The Journal of Applied Behavioral Science, 35,* 207–227. http://dx.doi.org/10.1177/0021886399352005

Fowler, S. M. (2006). Training across cultures: What intercultural trainers bring to diversity training. *International Journal of Intercultural Relations, 30,* 401–411. http://dx.doi.org/10.1016/j.ijintrel.2005.12.001

*Fradkin, H. R. (1980). *An exploratory study of the effects of specific training in counseling lesbian and gay clients on the counseling behaviors and attitudes of counselors-in-training* (Unpublished dissertation). University of North Carolina, Chapel Hill, NC.

Gallman, C. (2006). *Beyond pathologizing: Improving therapeutic interventions for African Americans* (Unpublished dissertation). Widener University, Chester, PA.

*Gannon, M. J., & Poon, J. M. L. (1997). Effects of alternative instructional approaches on cross-cultural training outcomes. *International Journal of Intercultural Relations, 21,* 429–446. http://dx.doi.org/10.1016/S0147-1767(96)00035-1

*Gany, F., & Thiel de Bocanegra, H. (1996). Maternal-child immigrant health training: Changing knowledge and attitudes to improve health care delivery. *Patient Education and Counseling, 27,* 23–31. http://dx.doi.org/10.1016/0738-3991(95)00786-5

*Garcia, J. A. (1996). *In search of cross-cultural competency: An empirical examination of cross-cultural training efforts in social work education* (Unpublished dissertation). University of Southern California, Los Angeles, CA.

Garvin, D. A., Cullen, P., & Datar, S. M. (2010). *Rethinking the MBA: Business education at a crossroads*. Watertown, MA: Harvard Business School Press Books.

*Gerla, J. J. (1999). *Panel presentations: A possible intervention for homophobia* (Unpublished dissertation). Stephen F. Austin State University, Nacogdoches, TX.

Geyskens, I., Krishnan, R., Steenkamp, J. E. M., & Cunha, P. V. (2009). A review and evaluation of meta-analysis practices in management research. *Journal of Management, 35,* 393–419. http://dx.doi.org/10.1177/0149206308328501

*Gharib, A., & Phillips, W. (2012). Assigning culture: An example of a cross-cultural assignment for teaching introductory psychology. *Psychology Learning & Teaching, 11,* 428–432. http://dx.doi.org/10.2304/plat.2012.11.3.428

Giangreco, A., Carugati, A., Sebastiano, A., & Bella, D. D. (2010). Trainees' reactions to training: Shaping groups and courses for happier trainees. *Journal of Human Resource Management, 21,* 2468–2487.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Gilbert, J., & Ivancevich, J. (2000). Valuing diversity: A tale of two organizations. *The Academy of Management Executive, 14,* 93–105.

*Glaser, T. S. (2006). *Learning to look deeply: How mindfulness meditation can help counselors overcome bias in the counseling process* (Unpublished dissertation). University of Minnesota, Minneapolis, MN.

*Goldberg, R. (1982). Attitude change among college students toward homosexuality. *College Health, 30,* 260–268.

*Goldstein, D. L., & Smith, D. H. (1999). The analysis of the effects of experiential training on Sojourners' cross-cultural adaptability. *International Journal of Intercultural Relations, 23,* 157–173. http://dx.doi.org/10.1016/S0147-1767(98)00030-3

Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation*. Mason, OH: Cengage/Wadsworth.

*Gorton, G. R., Sr. (1981). *Change in human sexuality attitudes* (Unpublished dissertation). University of Utah, Salt Lake City, UT.

*Govern, D. (1997). The effect of diversity awareness training on oral presentation ratings. *Dissertation Abstracts International, 58*(10-B), 5681.

*Grant, L. F. (2003). *Developing the cultural competence of registered nurses through nursing education* (Unpublished dissertation). University of Mississippi, Oxford, MS.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27. http://dx.doi.org/10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480. http://dx.doi.org/10.1037/0022-3514.74.6.1464

*Griswold, K., Kernan, J. B., Servoss, T. J., Saad, F. G., Wagner, C. M., & Zayas, L. E. (2006). Refugees and medical student training: Results of a programme in primary care. *Medical Education, 40,* 697–703. http://dx.doi.org/10.1111/j.1365-2929.2006.02514.x

Gunsch, D. (1993). Games augment diversity training. *The Personnel Journal, 72,* 78–83.

*Gursimsek, I. (2010). Sexual education and teacher candidates' attitudes towards sexuality. *Australian Journal of Guidance & Counselling, 20,* 81–90. http://dx.doi.org/10.1375/ajgc.20.1.81

*Guth, L. J., Lopez, D. F., Rojas, J., Clements, K. D., & Tyler, J. M. (2004). Experiential versus rational training: A comparison of student attitudes toward homosexuality. *Journal of Homosexuality, 48,* 83–102. http://dx.doi.org/10.1300/J082v48n02_05

*Guy-Walls, P. (2007). Exploring cultural competence practice in undergraduate social work education. *Education, 127,* 569–580.

Haines, E. L., & Sumner, K. E. (2006). Implicit measurement of attitudes, stereotypes, and self-concepts in organizations: Teaching old dogmas new tricks. *Organizational Research Methods, 9,* 536–553. http://dx.doi.org/10.1177/1094428106286540

*Hammer, M. R., & Martin, J. N. (1992). The effects of cross-cultural training on American managers in a Japanese-American Joint Venture. *Journal of Applied Communication Research, 20,* 161–182. http://dx.doi.org/10.1080/00909889209365327

*Handler, C. A. (1999). *The influence of climate for belief in the overseas mission on cross-cultural training effectiveness* (Unpublished dissertation). Louisiana State University, Baton Rouge, LA.

*Hanover, J., & Cellar, D. (1998). Environmental factors and the effectiveness of workforce diversity training. *Human Resource Development Quarterly, 9,* 105–124. http://dx.doi.org/10.1002/hrdq.3920090203

*Harrison, J. K. (1992). Individual and combined effects of behavior modeling and the cultural assimilator in cross-cultural management training. *Journal of Applied Psychology, 77,* 952–962. http://dx.doi.org/10.1037/0021-9010.77.6.952

*Hauenstein, N. M. A., Findley, R. A., & McDonald, D. P. (2010). Using situational judgment tests to assess training effectiveness: Lessons learned evaluating military equal opportunity advisor trainees. *Military Psychology, 22,* 262–281. http://dx.doi.org/10.1080/08995605.2010.492679

*Hayes, S. C., Bissett, R., Roget, N., Padilla, M., Kohlenberg, B. S., Fisher, G., . . . Niccolls, R. (2004). The impact of acceptance and commitment training and multicultural training on the stigmatizing attitudes and professional burnout of substance abuse counselors. *Behavior Therapy, 35,* 821–835. http://dx.doi.org/10.1016/S0005-7894(04)80022-4

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

*Henderson-King, D., & Kaleta, A. (2000). Learning about social diversity: The undergraduate experience and intergroup tolerance. *The Journal of Higher Education, 71,* 142–162. http://dx.doi.org/10.2307/2649246

Hequet, M. (1991). Flip charts and greasepaint. *Training, 28,* 65–73.

*Hernandez, C. R., & Gonzalez, M. Z. (2008). Effects of intergenerational interaction on aging. *Educational Gerontology, 34,* 292–305. http://dx.doi.org/10.1080/03601270701883908

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558. http://dx.doi.org/10.1002/sim.1186

*Hill, M. E., & Augoustinos, M. (2001). Stereotype change and prejudice reduction: Short- and long-term evaluation of a cross-cultural awareness programme. *Journal of Community & Applied Social Psychology, 11,* 243–262. http://dx.doi.org/10.1002/casp.629

*Hilliard, A. (2011). *Determinants of level of cultural competence prior to and after implementation of the neighborhood in a clinical pathway program* (Unpublished dissertation). Southern Nazarene University, Bethany, OK.

*Hillman, J., & Martin, R. A. (2002). Lessons about gay and lesbian lives: A spaceship exercise. *Teaching of Psychology, 29,* 308–311. http://dx.doi.org/10.1207/S15328023TOP2904_12

*Ho, M.-J., Yao, G., Lee, K.-L., Beach, M. C., & Green, A. R. (2008). Cross-cultural medical education: Can patient-centered cultural competency training be effective in non-Western countries? *Medical Teacher, 30,* 719–721. http://dx.doi.org/10.1080/01421590802232842

*Hodson, G., Choma, B. L., & Costello, K. (2009). Experiencing alienation: Effects of a simulation intervention on attitudes toward homosexuals. *Journal of Experimental Social Psychology, 45,* 974–978. http://dx.doi.org/10.1016/j.jesp.2009.02.010

*Hoff, J. G. (2005). *Students' perceptions of the culture learning process during the study abroad experience* (Unpublished dissertation). University of Minnesota, Minneapolis, MN.

*Hogan, D. E., & Mallott, M. (2005). Changing racial prejudice through diversity education. *Journal of College Student Development, 46,* 115–125. http://dx.doi.org/10.1353/csd.2005.0015

Holladay, C. L., Knight, J. L., Paige, D. L., & Quinones, M. A. (2003). The influence of framing on attitudes toward diversity training. *Human Resource Development Quarterly, 14,* 245–263. http://dx.doi.org/10.1002/hrdq.1065

Holladay, C. L., & Quinones, M. A. (2005). Reactions to diversity training: An international comparison. *Human Resource Development Quarterly, 16,* 529–545. http://dx.doi.org/10.1002/hrdq.1154

*Holladay, C. L., & Quinones, M. A. (2008). The influence of training focus and trainer characteristics on diversity training effectiveness. *Academy of Management Learning & Education, 7,* 343–354. http://dx.doi.org/10.5465/AMLE.2008.34251672

Hollister, L., Day, N., & Jesaitis, P. (1993). Diversity programs: Key to competitiveness or just another fad? *Organization Development Journal, 11,* 49–59.

*Hood, J. N., Muller, H. J., & Seitz, P. (2001). Attitudes of Hispanics and Anglos surrounding a workforce diversity intervention. *Hispanic Journal of Behavioral Sciences, 23,* 444–458. http://dx.doi.org/10.1177/0739986301234006

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

*Hostager, T. J., & De Meuse, K. P. (2008). The effects of a diversity learning experience on positive and negative diversity perceptions. *Journal of Business and Psychology, 23,* 127–139. http://dx.doi.org/10.1007/s10869-008-9085-x

*Howe, S. (2001). *An investigation of the diversity sensitivity training program: What is the impact of participation?* Unpublished Dissertation, University of Georgia, Athens, GA.

*Hoyt, J. M. (1987). *An investigation of a simulation game and the effects of gaming on nurses aides' attitudes toward the elderly in long term care* (Unpublished dissertation). University of San Francisco, San Francisco, CA.

Huber, S. G. (2013). Multiple learning approaches in the professional development of school leaders—Theoretical perspectives and empirical findings on self-assessment and feedback. *Educational Management Administration & Leadership, 41,* 527–540. http://dx.doi.org/10.1177/1741143213485469

Huffcut, A. I., & Arthur, W. J. (1995). Development of a new outlier statistic for meta-analysis data. *Journal of Applied Psychology, 80,* 327–334. http://dx.doi.org/10.1037/0021-9010.80.2.327

*Hughes, K. H., & Hood, L. J. (2007). Teaching methods and an outcome tool for measuring cultural sensitivity in undergraduate nursing students. *Journal of Transcultural Nursing, 18,* 57–62. http://dx.doi.org/10.1177/1043659606294196

*Hughes-White, G. F. (1991). *The effects of the Jigsaw method on achievement, friendship, and co-operative motives in cross-cultural classrooms* (Unpublished dissertation). University of Regina, Regina, Saskatchewan, Canada.

*Hurtado, S., Mayhew, M. J., & Engberg, M. E. (2012). Diversity courses and students' moral reasoning: A model of predispositions and change. *Journal of Moral Education, 41,* 201–224. http://dx.doi.org/10.1080/03057240.2012.670931

*Hussey, H. D., & Bisconti, T. L. (2010). Interventions to reduce sexual minority stigma in sororities. *Journal of Homosexuality, 57,* 566–587. http://dx.doi.org/10.1080/00918361003609184

*Hussey, H. D., Fleck, B. K. B., & Warner, R. M. (2010). Reducing student prejudice in diversity-infused core psychology classes. *College Teaching, 58,* 85–92. http://dx.doi.org/10.1080/87567550903418560

*Hylton, M. E. (2006). Online versus classroom-based instruction: A comparative study of learning outcomes in a diversity course. *The Journal of Baccalaureate Social Work, 11,* 102–114.

*Ibrahim, F. A., & Herr, E. L. (1976). Modification of attitudes toward disability: Differential effect of two educational modes. *Rehabilitation Counseling Bulletin, 26,* 29–36.

*Israel, T. (1998). *Comparing counselor education training models for working with lesbian, gay, and bisexual Clients* (Unpublished dissertation). Arizona State University, Tucson, AZ.

Ivancevich, J. M., & Gilbert, J. A. (2000). Diversity management: Time for a new approach. *Public Personnel Management, 29,* 75–92. http://dx.doi.org/10.1177/009102600002900106

*Jefferson, M. R. (2001). *The influence of diversity training on the attitudes and perceptions of freshman students*. (Unpublished dissertation) South Orange, NJ: Seton Hall University.

*Jeffreys, M. R., & Dogan, E. (2012). Evaluating the influence of cultural competence education on students' transcultural self-efficacy perceptions. *Journal of Transcultural Nursing, 23,* 188–197. http://dx.doi.org/10.1177/1043659611423836

Jehn, K. A., & Bezrukova, K. (2010). The faultline activation process and the effects of activated faultlines on coalition formation, conflict, and group outcomes. *Organizational Behavior and Human Decision Processes, 112,* 24–42. http://dx.doi.org/10.1016/j.obhdp.2009.11.008

*Johnson, L. M., Antle, B. F., & Barbee, A. P. (2009). Addressing disproportionality and disparity in child welfare: Evaluation of an anti-racism training for community service providers. *Children and Youth Services Review, 31,* 688–696. http://dx.doi.org/10.1016/j.childyouth.2009.01.004

*Jones, J. H. (1991). *The development of an educational program and an analysis of its effectiveness in altering racial attitudes of mid-level managers in a public agency* (Unpublished dissertation). The University of North Carolina at Chapel Hill, Chapel Hill, NC.

*Jones, T. S. (2008). *Preparing special educators: Infusing multicultural educational practices and lesson planning in pre-student teaching fieldwork* (Unpublished dissertation). University of Nevada, Las Vegas, Las Vegas, NV.

*Juarez, J. A., Marvel, K., Brezinski, K. L., Glazner, C., Towbin, M. M., & Lawton, S. (2006). Bridging the gap: A curriculum to teach residents cultural humility. *Research in Education, 38,* 97–102.

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. http://dx.doi.org/10.1177/000312240607100404

Kalinoski, Z. T., Steele-Johnson, D., Peyton, E. J., Leas, K. A., Steinke, J., & Bowling, N. A. (2013). A meta-analytic evaluation of diversity training outcomes. *Journal of Organizational Behavior, 34,* 1076–1104. http://dx.doi.org/10.1002/job.1839

*Kamfer, L., & Venter, D. (1994). First evaluation of a stereotype reduction workshop. *South African Journal of Psychology, 24,* 13–20. http://dx.doi.org/10.1177/008124639402400102

Kaplan, D. M. (2006). Can diversity training discriminate? Backlash to lesbian, gay, and bisexual diversity initiatives. *Employee Responsibilities and Rights Journal, 18,* 61–72. http://dx.doi.org/10.1007/s10672-005-9005-4

Karp, H. B., & Sammour, H. (2000). Workforce diversity: Choices in diversity training programs and dealing with resistance to diversity. *College Student Journal, 34,* 451–458.

*Katz, J. H. (1977). The effects of a systematic training program on the attitudes and behaviors of white people. *International Journal of Intercultural Relations, 1,* 77–89. http://dx.doi.org/10.1016/0147-1767(77)90008-6

*Katz, J. H., & Ivey, A. (1977). White awareness: The frontier of racism awareness training. *The Personnel and Guidance Journal, 55,* 485–489. http://dx.doi.org/10.1002/j.2164-4918.1977.tb04332.x

Kawakami, K., Dovidio, J. F., & van Kamp, S. (2007). The impact of counterstereotypic training and related correction processes on the application of stereotypes. *Group Processes & Intergroup Relations, 10,* 139–156. http://dx.doi.org/10.1177/1368430207074725

*Keillor, R. (1999). *Empathy and intergroup relations: A study in cross-cultural relationship Building* (Unpublished dissertation). Arizona State University, Tucson, AZ.

*Keim, J., Warring, D. F., & Rau, R. (2001). Impact of multicultural training on school psychology and education students. *Journal of Instructional Psychology, 28,* 249–252.

Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology, 93,* 59–69. http://dx.doi.org/10.1037/0021-9010.93.1.59

Kellough, J. E., & Naff, K. C. (2004). Responding to a wake-up call: An examination of federal agency diversity management programs. *Administration & Society, 36,* 62–90. http://dx.doi.org/10.1177/0095399703257269

*Kennedy, S. A. (1995). *Changing attitudes toward lesbian women and gay men through self-confrontation* (Unpublished dissertation). University of North Dakota, Grand Forks, ND.

Kepes, S., Banks, G. C., McDaniel, M., & Whetzel, D. L. (2012). Publication bias in the organizational sciences. *Organizational Research Methods, 15,* 624–662. http://dx.doi.org/10.1177/1094428112452760

Kepes, S., McDaniel, M. A., Brannick, M. T., & Banks, G. C. (2013). Meta-analytic reviews in the organizational sciences: Two meta-analytic schools on the way to MARS (the meta-analytic reporting standards).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Journal of Business and Psychology, 28, 123–143. http://dx.doi.org/10.1007/s10869-013-9300-2

*Kernahan, C., & Davis, T. (2007). Changing perspective: How learning about racism influences student awareness and emotion. *Teaching of Psychology, 34,* 49–52. http://dx.doi.org/10.1177/009862830703400111

*Kernahan, C., & Davis, T. (2010). What are the long-term effects of learning about racism? *Teaching of Psychology, 37,* 41–45. http://dx.doi.org/10.1080/00986280903425748

*Kilmnick, D. (2006). *Heterosexist attitudes: Changes following contact with an openly gay instructor* (Unpublished dissertation). City University of New York, New York, NY.

King, E. B., Gulick, L. M. V., & Avery, D. R. (2010). The divide between diversity training and diversity education: Integrating best practices. *Journal of Management Education, 34,* 891–906. http://dx.doi.org/10.1177/1052562909348767

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs: Pt. 1. Reactions. *Journal of the American Society for Training and Development, 13,* 3–9.

Kirkpatrick, D. L. (1979). Techniques for evaluating training programs. *Training and Development Journal, 33*(6), 78–92.

Kirkpatrick, D. L. (1996). Great ideas revisited. *Training & Development, 50,* 54–59.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). San Francisco, CA: Berrett-Koehler.

*Kitchens-Stephens, E. (2005). *The perceptions of beginning counselor education candidates' multicultural knowledge and awareness when comparing candidates who have successfully completed instruction in multicultural counseling with candidates who have not completed instruction in multicultural counseling* (Unpublished dissertation). Duquesne University, Pittsburg, PA.

*Klak, T., & Martin, P. (2003). Do university-sponsored international cultural events help students to appreciate "difference"? *International Journal of Intercultural Relations, 27,* 445–465. http://dx.doi.org/10.1016/S0147-1767(03)00033-6

Kleeb, R. (1989). Mobil drills holes through the color barrier. *Business and Society Review, 70,* 55–57.

*Kohl, B. G. (2005). *Evaluating cultural competency in social work field education: Expanding the vision for agency and university collaboration* (Unpublished dissertation). New York University, New York, NY.

Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning & Education, 4,* 193–212. http://dx.doi.org/10.5465/AMLE.2005.17268566

*Kracht, D. (1998). *Diversity training among manufacturing companies: Reaction and learning in a for-profit and not-for-profit work environment* (Unpublished dissertation). University of Sarasota, Sarasota, FL.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78,* 311–328. http://dx.doi.org/10.1037/0021-9010.78.2.311

Kulik, C. T., Pepper, M. B., Roberson, L., & Parker, S. K. (2007). The rich get richer: Predicting participation in voluntary diversity training. *Journal of Organizational Behavior, 28,* 753–769. http://dx.doi.org/10.1002/job.444

Kulik, C. T., Perry, E., & Bourhis, A. (2000). Ironic evaluation processes: Effects on thought suppression on evaluations of older job applicants. *Journal of Organizational Behavior, 21,* 689–711. http://dx.doi.org/10.1002/1099-1379(200009)21:6%3C689::aid-job52%3E3.0.co;2-w

Kulik, C. T., & Roberson, L. (2008a). Common goals and golden opportunities: Evaluations of diversity education in academic and organizational settings. *Academy of Management Learning & Education, 7,* 309–331. http://dx.doi.org/10.5465/AMLE.2008.34251670

Kulik, C. T., & Roberson, L. (2008b). Diversity initiative effectiveness: What organizations can (and cannot) expect from diversity recruitment, diversity training, and formal mentoring programs. In A. Brief (Ed.), *Diversity at work* (pp. 265–317). London, United Kingdom: Cambridge Press. http://dx.doi.org/10.1017/CBO9780511753725.010

*Lal, A. (2010). *The impact of an educational offering on the cultural competence of acute care registered nurses in an urban Midwest hospital in the United States* (Unpublished dissertation). University of Wisconsin–Milwaukee, Milwaukee, WI.

*Landis, D., Brislin, R. W., & Hulgus, J. F. (1985). Attributional training versus contact in acculturative learning: A laboratory study. *Journal of Applied Social Psychology, 15,* 466–482. http://dx.doi.org/10.1111/j.1559-1816.1985.tb02266.x

*Landis, D., Day, H. R., McGrew, P. L., Thomas, J. A., & Miller, A. B. (1976). Can a black "culture assimilator" increase racial understanding? *Journal of Social Issues, 32,* 169–183. http://dx.doi.org/10.1111/j.1540-4560.1976.tb02501.x

Lau, D., & Murnighan, J. K. (1998). Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of Management Review, 23,* 325–340.

*Law, D. Y. (1998). *An evaluation of a cultural diversity training program* (Unpublished dissertation). Auburn University, Auburn, AL.

*Lee, C. A., Anderson, M. A., & Hill, P. D. (2006). Cultural sensitivity education for nurses: A pilot study. *Journal of Continuing Education in Nursing, 37,* 137–141. http://dx.doi.org/10.3928/00220124-20060301-05

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin, 137,* 834–855. http://dx.doi.org/10.1037/a0024244

*Liberman, B. E., Block, C. J., & Koch, S. M. (2011). Diversity trainer preconceptions: The effects of trainer race and gender on perceptions of diversity trainer effectiveness. *Basic and Applied Social Psychology, 33,* 279–293. http://dx.doi.org/10.1080/01973533.2011.589327

*Lichtenstein, D., Lindstrom, L., & Povenmire-Kirk, T. (2008). Promoting multicultural competence: Diversity training for transition professionals. *Journal for Vocational Special Needs Education, 30,* 3–15.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

*Livosky, M., Pettijohn, T. F., & Capo, J. R. (2011). Reducing sexist attitudes as a result of completing an undergraduate psychology of gender course. *Psychology and Education: An Interdisciplinary Journal, 48,* 56–65.

*LoboPrabhu, S., King, C., Albucher, R., & Liberzon, I. (2000). A cultural sensitivity workshop for psychiatry residents. *Academic Psychiatry, 24,* 77–84. http://dx.doi.org/10.1176/appi.ap.24.2.77

Locke, E. A., & Latham, G. P. (1990). *A theory of goal-setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

*Lopez-Humphreys, M. (2011). *Intergroup dialogue an evaluation of a pedagogical model for teaching cultural competence within a framework of social justice in social work programs* (Unpublished dissertation). City University of New York, New York, NY.

*Luger, L. (2011). Enhancing cultural competence in staff working with people with drug and alcohol problems–A multidimensional approach to evaluating the impact of education. *Social Work Education, 30,* 223–235. http://dx.doi.org/10.1080/02615479.2011.540398

MacDonald, H. (1993). Cashing in on affirmative action: The diversity industry. *New Republic, 209,* 22–25.

*Madera, J., Neal, J. A., & Dawson, M. (2011). A strategy for diversity training: Focusing on empathy in the workplace. *Journal of Hospitality & Tourism Research, 35,* 469–487. http://dx.doi.org/10.1177/1096348010382240

*Majumdar, B., Browne, G., Roberts, J., & Carpio, B. (2004). Effects of cultural sensitivity training on health care provider attitudes and patient outcomes. *Journal of Nursing Scholarship, 36,* 161–166. http://dx.doi.org/10.1111/j.1547-5069.2004.04029.x

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

*Majumdar, B., Keystone, J. S., & Cuttress, L. A. (1999). Cultural sensitivity training among foreign medical graduates. *Medical Education, 33,* 177–184. http://dx.doi.org/10.1046/j.1365-2923.1999.00291.x

*Mak, A. S., & Buckingham, K. (2007). Beyond communication courses: Are there benefits in adding skills-based Excell™ sociocultural training? *International Journal of Intercultural Relations, 31,* 277–291. http://dx.doi.org/10.1016/j.ijintrel.2006.02.002

*Manese, J. E., Wu, J. T., & Nepomuceno, C. A. (2001). The effect of training on multicultural counseling competencies: An exploratory study over a ten-year period. *Journal of Multicultural Counseling and Development, 29,* 31–40. http://dx.doi.org/10.1002/j.2161-1912.2001.tb00501.x

*Manis, A. A. (2008). *Promoting multicultural and advocacy competence in counselor Trainees: Testing a critical pedagogical intervention* (Unpublished dissertation). University of Virginia, Charlottesville, VA.

*Mapp, S. C., McFarland, P., & Newell, E. A. (2007). The effect of a short-term study abroad class on students' cross-cultural awareness. *The Journal of Baccalaureate Social Work, 13,* 39–51.

Marshall, C., & Rossman, G. B. (1989). *Designing qualitative research*. Newbury Park, CA: Sage.

*Martin, M. M. (2006). *Cognitive and affective outcomes of a cultural diversity in business course in higher education* (Unpublished dissertation). Columbia, MO: University of Missouri–Columbia.

*Martinez, M. Y. (1995). *The application of a psychologically based model for cross-cultural counseling training within rehabilitation counselor education* (Unpublished dissertation). University of Arizona, Tucson, AZ.

*Mausehund, J., Timm, S., & King, A. (1995). Diversity training: Effects of an intervention treatment on nonverbal awareness. *Business Communication Quarterly, 58,* 27–30. http://dx.doi.org/10.1177/108056999505800107

*May, R. S. M. (2010). *Teacher education and preparation: Attitudes, beliefs, and perceptions of preparedness about teaching the culturally diverse student* (Unpublished dissertation). Texas A&M University–Commerce, Commerce, TX.

McBurney, D. H., & White, T. L. (2004). *Research methods*. Belmont, CA: Wadsworth/Thomson Learning.

*McCleskey, K. E. (1991). *The effects of a specialized training program about homosexuality on the knowledge and attitudes of counselor trainees* (Unpublished dissertation). East Texas State University, Commerce, TX.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in America? *The Journal of Conflict Resolution, 25,* 563–579.

*McCool, M. A., Du Toit, F., Petty, C. R., & McCauley, C. (2006). The impact of a program of prejudice-reduction seminars in South Africa. *Journal of Applied Social Psychology, 36,* 586–613. http://dx.doi.org/10.1111/j.0021-9029.2006.00020.x

*Medina-Walpole, A., Mooney, C. J., Lyness, J. M., Lambert, D. R., & Lurie, S. J. (2012). Medical student attitudes toward patients in diverse care settings: The impact of a patient evaluation course. *Teaching and Learning in Medicine, 24,* 117–121. http://dx.doi.org/10.1080/10401334.2012.664533

*Middleton, V. A. (2002). Increasing preservice teachers' diversity beliefs and commitment. *The Urban Review, 34,* 343–361. http://dx.doi.org/10.1023/A:1021372801442

*Mio, J. S. (1989). Experiential involvement as an adjunct to teaching cultural sensitivity. *Journal of Multicultural Counseling and Development, 17,* 38–46. http://dx.doi.org/10.1002/j.2161-1912.1989.tb00414.x

*Moffat, J., & Tung, J.-Y. (2004). Evaluating the effectiveness of culture brokering training to enhance cultural competence of independent living center staff. *Journal of Vocational Rehabilitation, 20,* 59–69.

Mor-Barak, M. (2005). *Managing diversity: Toward a globally inclusive workplace*. Thousand Oaks, CA: Sage.

*Morin, S. F. (1974). Educational programs as a means of changing attitudes toward gay people. *Homosexual Counseling Journal, 1,* 160–165.

Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *The British Journal of Mathematical and Statistical Psychology, 53,* 17–29. http://dx.doi.org/10.1348/000711000159150

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105–125. http://dx.doi.org/10.1037/1082-989X.7.1.105

*Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M., & Handelsman, J. (2013). *Evidence for the efficacy of a diversity intervention with academic scientists*. Paper presented at the Society of Experimental Social Psychology, Berkeley, CA.

*Moyer, R. S., & Nath, A. (1998). Some effects of brief training interventions on perceptions of sexual harassment. *Journal of Applied Social Psychology, 28,* 333–356. http://dx.doi.org/10.1111/j.1559-1816.1998.tb01709.x

*Murphy, M., Park, J., & Lonsdale, N. (2006). Marriage and family therapy students' change in multicultural counseling competencies after a diversity course. *Contemporary Family Therapy: An International Journal, 28,* 303–311. http://dx.doi.org/10.1007/s10591-006-9009-8

*Mysore, A. R. (2004). *Attitudes of preservice teachers toward issues in multicultural education* (Unpublished dissertation). University of Arkansas, Fayetteville, AR.

National Urban League. (2009). *Diversity practices that work: The American worker speaks II*. Retrieved May 11, 2011, from http://www.nul.org/sites/default/files/Diversity_Practices_That_work_2009.pdf

*Neville, H. A., & Furlong, M. (1994). The impact of participation in a cultural awareness program on the racial attitudes and social behaviors of first-year college students. *Journal of College Student Development, 35,* 371–377.

*Neville, H. A., Heppner, M. J., Louie, C. E., Thompson, C. E., Brooks, L., & Baker, C. E. (1996). The impact of multicultural training on white racial identity attitudes and therapy competencies. *Professional Psychology, Research and Practice, 27,* 83–89. http://dx.doi.org/10.1037/0735-7028.27.1.83

*Nguyen, H., Tran, L. V., Derous, E., Lopez, C. M., & Siu, C. (2012). *Effectiveness of a multi-media perspective-taking training program on reducing discrimination*. Paper presented at the Society for Industrial–Organizational Psychology, San Diego, CA.

Noe, R. A. (2010). *Employee training and development*. New York, NY: McGraw-Hill/Irwin.

*Nokes, K. M., Nickitas, D. M., Keida, R., & Neville, S. (2005). Does service-learning increase cultural competency, critical thinking, and civic engagement? *The Journal of Nursing Education, 44,* 65–70.

Paluck, E. L. (2006). Diversity training and intergroup contact: A call to action research. *Journal of Social Issues, 62,* 577–595. http://dx.doi.org/10.1111/j.1540-4560.2006.00474.x

*Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology, 96,* 574–587. http://dx.doi.org/10.1037/a0011989

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology, 60,* 339–367. http://dx.doi.org/10.1146/annurev.psych.60.110707.163607

*Parker, W. M., Moore, M. A., & Neimeyer, G. J. (1998). Altering White racial identity and interracial comfort through multicultural training. *Journal of Counseling and Development, 76,* 302–310. http://dx.doi.org/10.1002/j.1556-6676.1998.tb02546.x

*Pedersen, A., & Barlow, F. K. (2008). Theory to social action: A university-based strategy targeting prejudice against Aboriginal Austra-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

lians. *Australian Psychologist, 43,* 148–159. http://dx.doi.org/10.1080/00050060802318587

*Pedersen, P. J. (2010). Assessing intercultural effectiveness outcomes in a year-long study abroad program. *International Journal of Intercultural Relations, 34,* 70–80. http://dx.doi.org/10.1016/j.ijintrel.2009.09.003

Pendry, L. F., Driscoll, D. M., & Field, C. T. (2007). Diversity training: Putting theory into practice. *Journal of Occupational and Organizational Psychology, 80,* 227–250. http://dx.doi.org/10.1348/096317906X118397

*Perry, E. L., Kulik, C. T., & Schmidtke, J. M. (1998). Individual differences in the effectiveness of sexual harassment awareness training. *Journal of Applied Social Psychology, 28,* 698–723. http://dx.doi.org/10.1111/j.1559-1816.1998.tb01727.x

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology, 49,* 65–85. http://dx.doi.org/10.1146/annurev.psych.49.1.65

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

*Pettijohn, T., & Walzer, A. (2008). Reducing racism, sexism, and homophobia in college students by completing a psychology of prejudice course. *College Student Journal, 42,* 459–468.

Pew Research Center. (2012). *Generation gap at the polls is echoed in attitudes on entitlements*. Retrieved December 20, 2012, from www.pewsocialtrends.org

*Pilkington, W. J. (1993). *Effects of a communication program on nursing home staff's attitudes and communication skills and residents' levels of life satisfaction* (Unpublished dissertation). St. John's University, Queens, NY.

*Preusser, M. K., Bartels, L. K., & Nordstrom, C. R. (2011). Sexual harassment training: Person versus machine. *Public Personnel Management, 40,* 47–62. http://dx.doi.org/10.1177/009102601104000104

*Probst, T. M. (2003). Changing attitudes over time: Assessing the effectiveness of a workplace diversity course. *Teaching of Psychology in the Schools, 30,* 236–239. http://dx.doi.org/10.1207/S15328023TOP3003_09

Quinones, M. A. (1997). Contextual influences on training effectiveness. In M. A. Quinones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace: applications of psychological research* (pp. 177–199). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10260-007

Raelin, J. A. (1997). A model of work-based learning. *Organization Science, 8,* 563–578. http://dx.doi.org/10.1287/orsc.8.6.563

*Randolph, G., Landis, D., & Tzeng, O. C. S. (1977). The effects of time and practice upon culture assimilator training. *International Journal of Intercultural Relations, 1,* 105–119. http://dx.doi.org/10.1016/0147-1767(77)90034-7

*Rehg, M. T., Gundalch, M. J., & Grigorian, R. A. (2012). Examining the influence of cross-cultural training on cultural intelligence and specific self-efficacy. *Cross Cultural Management, 19,* 215–232. http://dx.doi.org/10.1108/13527601211219892

*Reinhardt, B. M. (1994). *Effects of gay and lesbian speaker panels on self-report measures of individual homophobia* (Unpublished dissertation). Texas A&M University, College Station, TX.

Remer, P., & Remer, R. (2000). The alien invasion exercise: Creating an experience of diversity. *International Journal of Action Methods: Psychodrama, Skill Training, & Role Playing, 52,* 147–154.

*Remer, R. R. (2008). *Influence of diversity courses on undergraduates' ethnocultural empathy, openness to diversity, and awareness of privilege and oppression within a mastery or performance classroom context* (Unpublished dissertation). Lexington, KY: University of Kentucky.

*Remmert, A. A. (1993). *The impact of multicultural in-service education on the cross-cultural adaptability of public school teachers* (Unpublished dissertation). Florida State University, Tallahassee, FL.

*Reynolds, L. (2010). Aging and disability awareness training for drivers of a metropolitan taxi company. *Activities, Adaptation and Aging, 34,* 17–29. http://dx.doi.org/10.1080/01924780903552279

*Rhyne, D. C. (1973). Adult education and the rational-irrational dimension of prejudice. *Adult Education, 23,* 203–233. http://dx.doi.org/10.1177/074171367302300303

*Riggs, A. D., Rosenthal, A. R., & Smith-Bonahue, T. (2011). The impact of a combined cognitive–affective intervention on pre-service teachers' attitudes, knowledge, and anticipated professional behaviors regarding homosexuality and gay and lesbian issues. *Teaching and Teacher Education, 27,* 201–209. http://dx.doi.org/10.1016/j.tate.2010.08.002

Robb, L. A., & Doverspike, D. (2001). Self-reported proclivity to harass as a moderator of the effectiveness of sexual harassment-prevention training. *Psychological Reports, 88,* 85–88. http://dx.doi.org/10.2466/pr0.2001.88.1.85

Roberson, L., Kulik, C. T., & Pepper, M. B. (2001). Designing effective diversity training: Influence of group composition and trainee experience. *Journal of Organizational Behavior, 22,* 871–885. http://dx.doi.org/10.1002/job.117

*Roberson, L., Kulik, C. T., & Pepper, M. B. (2002). Assessing instructor cultural competence in the classroom: An instrument and a development process. *Journal of Management Education, 26,* 40–55. http://dx.doi.org/10.1177/105256290202600104

Roberson, L., Kulik, C. T., & Pepper, M. B. (2003). Using needs assessment to resolve controversies in diversity training design. *Group & Organization Management, 28,* 148–174. http://dx.doi.org/10.1177/1059601102250028

Roberson, L., Kulik, C. T., & Pepper, M. B. (2009). Individual and environmental factors influencing the use of transfer strategies after diversity training. *Group & Organization Management, 34,* 67–89. http://dx.doi.org/10.1177/1059601108329732

*Robinson, B., & Bradley, L. J. (1997). Multicultural training for undergraduates: Developing knowledge and awareness. *Journal of Multicultural Counseling and Development, 25,* 281–289. http://dx.doi.org/10.1002/j.2161-1912.1997.tb00349.x

*Rogers-Sirin, L., & Sirin, S. R. (2009). Cultural competence as an ethical requirement: Introducing a new educational model. *Journal of Diversity in Higher Education, 2,* 19–29. http://dx.doi.org/10.1037/a0013762

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781412984997

*Rouh, P. A. (2001). *The effects of diversity training on recognizing gender differences in a corporate environment* (Unpublished dissertation). University of North Texas, Denton, TX.

*Rousey, S. L. (2010). *The effects of viewing film clips on multicultural awareness and empathy in pre-service counselors* (Unpublished dissertation). University of North Carolina at Charlotte, Charlotte, NC.

*Rowell, P. C., & Benshoff, J. M. (2008). Using personal growth groups in multicultural counseling courses to foster students' ethnic identity development. *Counselor Education and Supervision, 48,* 2–15. http://dx.doi.org/10.1002/j.1556-6978.2008.tb00058.x

*Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81,* 856–868. http://dx.doi.org/10.1037/0022-3514.81.5.856

*Rudolph, J. (1989). Effects of a workshop on mental health practitioners' attitudes toward homosexuality and counseling effectiveness. *Journal of Counseling and Development, 68,* 81–85. http://dx.doi.org/10.1002/j.1556-6676.1989.tb02499.x

Rynes, S., & Rosen, B. (1994). What makes diversity programs work? *HR Magazine, 39,* 67–68.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Rynes, S., & Rosen, B. (1995). A field survey of factors affecting the adoption and perceived success of diversity training. *Personnel Psychology, 48,* 247–270. http://dx.doi.org/10.1111/j.1744-6570.1995.tb01756.x

*Sakurai, T., McCall-Wolf, F., & Kashima, E. S. (2010). Building intercultural links: The impact of a multicultural intervention programme on social ties of international students in Australia. *International Journal of Intercultural Relations, 34,* 176–185. http://dx.doi.org/10.1016/j.ijintrel.2009.11.002

Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology, 52,* 471–499. http://dx.doi.org/10.1146/annurev.psych.52.1.471

*Sanchez, J. I., & Medkik, N. (2004). The effects of diversity awareness training on differential treatment. *Group & Organization Management, 29,* 517–536. http://dx.doi.org/10.1177/1059601103257426

*Sanchez-Burks, J., Lee, F., Nisbett, R., & Ybarra, O. (2007). Cultural training based on a theory of relational ideology. *Basic and Applied Social Psychology, 29,* 257–268. http://dx.doi.org/10.1080/01973530701503184

*Sanner, S., Baldwin, D., Cannella, K. A. S., Charles, J., & Parker, L. (2010). The impact of cultural diversity forum on students' openness to diversity. *Journal of Cultural Diversity, 17,* 56–61.

*Scher, L. J. (2008). *Beyond acceptance: An evaluation of the safe zone project in a clinical psychology doctoral program* (Unpublished dissertation). Brookville, NY: Long Island University.

Schwartz, J. (2010). Bullying, suicide, punishment: A case of a Rutgers student's suicide. *New York Times*. Retrieved from www.nytimes.com/2010/10/3/weekinreview03/schwartz.html?_r=0

*Seguin, C. A. K. (2002). *Evaluating educational sustainability: Effects of training and program delivery on adults implementing diversity awareness programs for children at 11 Michigan summer camps* (Unpublished dissertation). Wayne State University, Detroit, MI.

*Shergill, A. S. (1997). *An evaluation of the social cultural competency for success training program for the acquisition of intercultural interpersonal competency skills among health care trainees* (Unpublished dissertation). University of British Columbia, Vancouver, British Columbia, Canada.

*Shields, S. A., Zawadzki, M. J., & Johnson, R. N. (2011). The impact of the Workshop Activity for Gender Equity Simulation in the Academy (WAGES–Academic) in demonstrating cumulative effects of gender bias. *Journal of Diversity in Higher Education, 4,* 120–129. http://dx.doi.org/10.1037/a0022953

Sitzman, T., Ely, K., Brown, K., & Bauer, K. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9,* 169–191. http://dx.doi.org/10.5465/AMLE.2010.51428542

Smith, T., Constantine, M., Dunn, T., Dinehart, J., & Montoya, J. (2006). Multicultural education in the mental health professions: A meta-analytic review. *Journal of Counseling Psychology, 53,* 132–145. http://dx.doi.org/10.1037/0022-0167.53.1.132

*Soble, J. R., Spanierman, L. B., & Liao, H.-Y. (2011). Effects of a brief video intervention on White university students' racial attitudes. *Journal of Counseling Psychology, 58,* 151–157. http://dx.doi.org/10.1037/a0021158

Spell, C., & Blum, T. (2005). The adoption of workplace substance abuse programs: Strategic choice and institutional perspectives. *Academy of Management Journal, 48,* 1125–1142. http://dx.doi.org/10.5465/AMJ.2005.19573113

*St. Clair, A., & McKenry, L. (1999). Preparing culturally competent practitioners. *The Journal of Nursing Education, 38,* 228–234.

*Stebbins, C. A. (2005). *Enhancing empathy in medical students using FLEX Care™ communication training* (Unpublished dissertation). Iowa State University, Ames, IA.

*Steed, R. (2010). Attitudes and beliefs of occupational therapists participating in a cultural competency workshop. *Occupational Therapy International, 17,* 142–151. http://dx.doi.org/10.1002/oti.299

Steel, P. D., & Kammeyer-Mueller, J. D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology, 87,* 96–111. http://dx.doi.org/10.1037/0021-9010.87.1.96

*Steinfeldt, J. A., & Wong, Y. J. (2010). Multicultural training on American Indian issues: Testing the effectiveness of an intervention to change attitudes toward Native-themed mascots. *Cultural Diversity & Ethnic Minority Psychology, 16,* 110–115. http://dx.doi.org/10.1037/a0018633

*Stella, C. S. C., Forlin, C., & Lan, A. M. (2007). The influence of an inclusive education course on attitude change of pre-service secondary teachers in Hong Kong. *Asia-Pacific Journal of Teacher Education, 35,* 161–179. http://dx.doi.org/10.1080/13598660701268585

*Stewart, T. L., LaDuke, J. R., Bracht, C., Sweet, B. A. M., & Gamarel, K. E. (2003). Do the "eyes" have it? A program evaluation of Jane Elliott's "Blue-Eyes/Brown-Eyes" diversity training exercise. *Journal of Applied Social Psychology, 33,* 1898–1921. http://dx.doi.org/10.1111/j.1559-1816.2003.tb02086.x

*Stewart, T. L., Latu, I. M., Kawakami, K., & Myers, A. C. (2010). Consider the situation: Reducing automatic stereotyping through situational attribution training. *Journal of Experimental Social Psychology, 46,* 221–225. http://dx.doi.org/10.1016/j.jesp.2009.09.004

*Stowe, A. M. (2002). *An examination of two approaches to experiential learning and the impact of service-learning on multicultural awareness* (Unpublished dissertation). Auburn University, Auburn, AL.

Stratton, J., Canales, C., Armas, A., & Miller, N. (2006). Positive stereotyping: Influence tactic for prejudice reduction? *Social Influence, 1,* 265–287. http://dx.doi.org/10.1080/15534510601008882

Stroup, C. (1998). *Diversity training and gender communication dynamics* (Unpublished master's thesis). Pennsylvania State University, University Park, PA.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7–24). Chicago, IL: Nelson-Hall.

*Tang, T. S., Fantone, J. C., Bozynski, M. E. A., & Adams, B. S. (2002). Implementation and evaluation of an undergraduate sociocultural medicine program. *Academic Medicine, 77,* 578–585. http://dx.doi.org/10.1097/00001888-200206000-00019

*Tang, T. S., Hernandez, E. J., & Adams, B. S. (2004). "Learning by teaching": A peer-teaching model for diversity training in medical school. *Teaching and Learning in Medicine, 16,* 60–63. http://dx.doi.org/10.1207/s15328015tlm1601_12

*Taras, V., Caprar, D. V., Rottig, D., Sarala, R. M., Zarkaria, N., Zhao, F., . . . Huang, V. Z. (2013). A global classroom? Evaluating the effectiveness of global virtual collaboration as a teaching tool in management education. *Academy of Management Learning & Education, 12,* 414–435. http://dx.doi.org/10.5465/amle.2012.0195

*Taton, J. L. (2008). *The impact of an instructional unit regarding gay and lesbian issues on knowledge and comfort level of course participants* (Unpublished dissertation). Central Michigan University, Mount Pleasant, MI.

Taylor, P. J., Russ-Eft, D. F., & Chan, D. W. L. (2005). A meta-analytic review of behavior modeling training. *Journal of Applied Psychology, 90,* 692–709. http://dx.doi.org/10.1037/0021-9010.90.4.692

*Taylor-Ritzler, T., Balcazar, F., Dimpfl, S., Suarez-Balcazar, Y., Willis, C., & Schiff, R. (2008). Cultural competence training with organizations serving people with disabilities from diverse cultural backgrounds. *Journal of Vocational Rehabilitation, 29,* 77–91.

Tesluk, P. E., Farr, J. L., Mathieu, J. E., & Vance, R. J. (1995). Generalization of employee involvement training to the job setting: Individual and situational effects. *Personnel Psychology, 48,* 607–632. http://dx.doi.org/10.1111/j.1744-6570.1995.tb01773.x

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Thatcher, S. M. B., & Patel, P. C. (2012). Group faultlines: A review, integration, and guide to future research. *Journal of Management, 38,* 969–1009. http://dx.doi.org/10.1177/0149206311426187

*Theis-Cole, D. E. (1995). *The effectiveness of a prejudice reduction workshop among teacher education students* (Unpublished dissertation). Texas A&M University, College Station, TX.

*Thomas, C. H. (2008). *An analysis of the effects of diversity training on employment practices* (Unpublished dissertation). University of Phoenix, Phoenix, AZ.

Thomas, R. R. (1996). Managing diversity: A conceptual framework. In S. E. Jackson & Associates. (Eds.), *Diversity in the workplace: Human resources initiatives* (pp. 306–319). New York, NY: Guilford Press.

Thomas, R. R. (1999). Diversity management: Some measurement criteria. *Employment Relations Today, 25,* 49–62. http://dx.doi.org/10.1002/ert.3910250406

*Thomas, V. J., & Cohn, T. (2006). Communication skills and cultural awareness courses for healthcare professionals who care for patients with sickle cell disease. *Journal of Advanced Nursing, 53,* 480–488. http://dx.doi.org/10.1111/j.1365-2648.2006.03741.x

*Timm, S. A. (1997). *A comparative analysis of the influence of listening/nonverbal communication training on multicultural sensitivity* (Unpublished dissertation). Northern Illinois University, DeKalb, IL.

*Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology, 100,* 1027–1042. http://dx.doi.org/10.1037/a0022308

*Tolleson Knee, R. (1999). *Service-learning in social work education: Building democracy through informed citizenship* (Unpublished dissertation). University of Denver, Denver, CO.

*Tran, M. T., Young, R. L., & Di Lella, J. D. (1994). Multicultural education courses and the student teacher: Eliminating stereotypical attitudes in our ethnically diverse classroom. *Journal of Teacher Education, 45,* 183–189. http://dx.doi.org/10.1177/0022487194045003004

*Turner, S. M. (1986). *Development and evaluation of a racial awareness curriculum* (Unpublished dissertation). Cornell University, Ithaca, NY.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139,* 352–402. http://dx.doi.org/10.1037/a0028446

*Van Soest, D. J. (1996). Impact of social work education on student attitudes and behavior concerning oppression. *Journal of Social Work Education, 32,* 191–202.

*VonDras, D. D., & Lor-Vang, M. N. (2004). Using an Internet activity to enhance students' awareness of age bias in social perceptions. *Educational Gerontology, 30,* 261–273. http://dx.doi.org/10.1080/03601270490275572

Waddock, S., & Lozano, J. M. (2013). Developing more holistic management education: Lessons learned from two programs. *Academy of Management Learning & Education, 12,* 265–284. http://dx.doi.org/10.5465/amle.2012.0002

*Wade, P., & Bernstein, B. L. (1991). Cultural sensitivity training and counselor's race: Effects on black female clients' perceptions and attrition. *Journal of Counseling Psychology, 38,* 9–15. http://dx.doi.org/10.1037/0022-0167.38.1.9

*Waight, D., & Madera, J. M. (2011). Diversity training: Examining minority employees' organizational attitudes. *Worldwide Hospitality and Tourism Themes, 3,* 365–376. http://dx.doi.org/10.1108/17554211111162471

*Wang, V. O. (1998). Curriculum evaluation and assessment of multicultural genetic counselor education. *Journal of Genetic Counseling, 7,* 87–111. http://dx.doi.org/10.1023/A:1022820229328

*Warring, D., Keim, J., & Rau, R. (1998). Multicultural training for students and its impact. *Action in Teacher Education, 20,* 56–63. http://dx.doi.org/10.1080/01626620.1998.10462917

*Weathersby, A. (2004). *The effects of multicultural education on White racial consciousness* (Unpublished dissertation). Washington State University, Pullman, WA.

Werner, J. M., & DeSimone, R. L. (2009). *Human resource development*. Mason, OH: Cengage/South-Western.

Wheeler, M. (1998). Measuring diversity: A strategy for organizational effectiveness. *Employment Relations Today, 25,* 61–68. http://dx.doi.org/10.1002/ert.3910250107

White, M. B. (1998). Workshop equals profit measuring change at Sears. *Diversity Factor, 7,* 7–10.

*Wideman, R. A. E. (2005). *Empathy development in undergraduate students through the cross-cultural learning experience* (Unpublished dissertation). Regent University, Virginia Beach, VA.

Wiethoff, C. (2004). Motivation to learn and diversity training: Application of the theory of planned behavior. *Human Resource Development Quarterly, 15,* 263–278. http://dx.doi.org/10.1002/hrdq.1103

*Wiggins, R. A., Follo, E. J., & Eberly, M. B. (2007). The impact of a field immersion program on pre-service teachers' attitudes toward teaching in culturally diverse classrooms. *Teaching and Teacher Education, 23,* 653–663. http://dx.doi.org/10.1016/j.tate.2007.02.007

*Williams, C. C. (2005). Training for cultural competence: Individual and group processes. *Journal of Ethnic & Cultural Diversity in Social Work, 14,* 111–143. http://dx.doi.org/10.1300/J051v14n01_06

Wilson, D. B. (2009). Systematic coding. In H. Cooper, L. Hedges, & J. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 159–176). New York, NY: Russell Sage Foundation.

Winne, P. (2005). Key issues in modeling and applying research on self-regulated learning. *Applied Psychology, 54,* 232–238. http://dx.doi.org/10.1111/j.1464-0597.2005.00206.x

Winterle, M. J. (1993). Toward diversity, with carrots and sticks. *Across the Board, 30,* 50.

*Worchel, S., & Mitchell, T. R. (1972). An evaluation of the effectiveness of the culture assimilator in Thailand and Greece. *Journal of Applied Psychology, 56,* 472–479. http://dx.doi.org/10.1037/h0033759

*Wortman, T. I. (2002). *Psychosocial effects of studying abroad: Openness to diversity* (Unpublished dissertation). Pennsylvania State University, University Park, PA.

*Zarubin, A. (2008). *Effect of attitudes on results of diversity training* (Unpublished master's thesis). Bowling Green State University, Bowling Green, OH.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A

Brief Descriptions and Statistics of Included Studies

Author, year	*N*	Training context			Training design				Methodological moderators study rigor	*g* (*SE*)
		Setting	Training approach	Training attendance	Training focus	Training duration	Training type	Instruction		
Abernethy, 2005	15	W	Stand-alone	Voluntary	One	4 days	Both	Many	Observational	.84 (.29)
Aldridge, 2013	56	W	Stand-alone	Voluntary	One	3 month	Both	Many	Quasi-experimental	.67 (.27)
Alonso, 2005										
Sample a	42	E	Stand-alone	Voluntary	One	Semester	Both	Many	Observational	.83 (.29)
Sample b	42	E	Stand-alone	Voluntary	One	Semester	Both	Many	Observational	.72 (.21)
Altshuler, Sussman, & Kachur, 2003										
Training 1	24	E	Stand-alone	Voluntary	One	3 hr	Both	Many	Observational	.00 (.29)
Training 2	24	E	Stand-alone	Voluntary	One	3 hr	Both	Many	Observational	−.19 (.33)
Amatea, Cholewa, & Mixon, 2012	138	E	Stand-alone	Voluntary	One	Semester	Awareness	Many	Observational	.88 (.10)
Anderson, 1981	64	E	Stand-alone	Voluntary	One	2 hr	Awareness	One	Quasi-experimental	.73 (.27)
Arizaga, 1999	55	E	Stand-alone	Voluntary	One	8 hr	Both	Many	Experimental	.37 (.27)
Armour, Bain, & Rubio, 2004	11	W	Stand-alone	Voluntary	Multiple	18 hr	Both	Many	Observational	.88 (.34)
Baba & Hebert, 2004	170	W	Stand-alone	Mandatory	One	8 weeks	Awareness	Many	Observational	−.18 (.08)
Bailey, Barr, & Bunting, 2001	57	W	Stand-alone	Voluntary	One	Unknown	Both	Many	Quasi-experimental	.47 (.26)
Baker, 2009	177	E	Stand-alone	Voluntary	One	20 min	Awareness	One	Quasi-experimental	.00 (.15)
Ballou, 1996	308	E	Stand-alone	Voluntary	One	Semester	Awareness	One	Quasi-experimental	.23 (.11)
Barker & Hartel, 2004	62	E	Stand-alone	Voluntary	One	6 weeks	Awareness	One	Experimental	.77 (.26)
Bauer, McAulife, Nay, & Chenco, 2013	112	W	Stand-alone	Voluntary	One	3 hr	Awareness	Many	Observational	−.46 (.14)
Bhawuk, 1998	102	E	Stand-alone	Voluntary	One	Overnight	Awareness	One	Experimental	.78 (.29)
Blakely, Blakely, & Moorman, 1998	176	E	Stand-alone	Voluntary	One	Unknown	Both	One	Quasi-experimental	.26 (.16)
Bloch, 2012	21	E	Stand-alone	Voluntary	One	4 hr	Both	Many	Quasi-experimental	.19 (.23)
Boswell, 2012	43	E	Stand-alone	Voluntary	One	Semester	Awareness	Many	Observational	.53 (.16)
Brathwaite & Majumdar, 2006	76	W	Stand-alone	Mandatory	One	5 weeks	Both	Many	Observational	.27 (.12)
Brown, 2004	100	E	Stand-alone	Mandatory	Multiple	10 weeks	Awareness	Many	Quasi-experimental	−.13 (.20)
Brown, Parham, & Yonker, 1996	35	E	Stand-alone	Voluntary	One	16 weeks	Both	Many	Observational	.57 (.37)
Burris, 1992	67	W	Stand-alone	Voluntary	One	3 hr	Both	One	Quasi-experimental	1.83 (.29)
Bush & Ingram, 2001	122	W	Stand-alone	Voluntary	Multiple	2 hr	Behavior	One	Observational	.37 (.11)
Butler et al., 2012	80	E	Stand-alone	Voluntary	Multiple	Unknown	Behavior	One	Quasi-experimental	.40 (.23)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A (*continued*)

Author, year	*N*	Training context			Training design				Methodological moderators study rigor	*g* (*SE*)
		Setting	Training approach	Training attendance	Training focus	Training duration	Training type	Instruction		
Byington, Fischer, Walker, & Freedman, 1997	48	W	Stand-alone	Voluntary	Inclusive	2 weeks	Awareness	One	Observational	.51 (.16)
Caffrey et al., 2005	7	E	Integrated	Mandatory	One	2 years	Both	Many	Quasi-experimental	1.25 (.46)
Cap, 1995	82	E	Stand-alone	Voluntary	One	Unknown	Awareness	One	Experimental	.69 (.28)
Carpenter, 1997	78	E	Stand-alone	Voluntary	One	80 hr	Behaviors	Many	Quasi-experimental	1.24 (.28)
Carrell, 1997										
Sample a	237	E	Stand-alone	Mandatory	Inclusive	2 weeks	Awareness	One	Experimental	−.15 (.25)
Sample b	237	E	Stand-alone	Mandatory	Inclusive	2 weeks	Awareness	One	Experimental	.50 (.22)
Sample c	237	E	Stand-alone	Mandatory	Inclusive	2 weeks	Awareness	One	Experimental	.25 (.35)
Sample d	237	E	Integrated	Voluntary	Inclusive	Semester	Both	Many	Experimental	.88 (.27)
Carter et al., 2006	196	E	Integrated	Mandatory	One	3 hr	Both	Many	Observational	.02 (.07)
Cascio & Bass, 1976	2292	E	Stand-alone	Mandatory	One	4 hr	Awareness	One	Observational	.04 (.02)
Case, 2007	147	E	Stand-alone	Voluntary	One	15 weeks	Awareness	One	Observational	.03 (.08)
Case & Stewart, 2010	108	E	Stand-alone	Voluntary	One	Semester	Awareness	One	Quasi-experimental	.42 (.22)
Case & Stewart, 2013	132	E	Stand-alone	Voluntary	One	Unknown	Awareness	One	Observational	.41 (.09)
Castillo et al., 2007	84	E	Integrated	Voluntary	One	15 weeks	Awareness	Many	Quasi-experimental	.29 (.22)
Cates, 2006	47	E	Stand-alone	Voluntary	One	90 hr	Both	One	Quasi-experimental	.96 (.30)
Celik et al., 2012	31	W	Stand-alone	Unknown	Multiple	16 hr	Both	Many	Observational	1.26 (.52)
Chang, 2002	186	E	Stand-alone	Mandatory	One	14 weeks	Awareness	Many	Quasi-experimental	.17 (.16)
Chrobot-Mason, 2004	53	E	Stand-alone	Voluntary	One	16 weeks	Both	Many	Quasi-experimental	.42 (.28)
Clark, 1998	193	E	Stand-alone	Voluntary	One	15 weeks	Both	Many	Quasi-experimental	.18 (.16)
Clevenger, 2011	74	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	.37 (.12)
Clinton, 1983	76	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	.36 (.31)
Cole, Case, Rios, & Curtin, 2011	173	E	Stand-alone	Voluntary	Multiple	Semester	Awareness	One	Quasi-experimental	.11 (.16)
Colvin-Burque, Davis-Maye, & Zugazaga, 2007	110	E	Stand-alone	Unknown	Multiple	Semester	Awareness	Many	Observational	.47 (.10)
Combs & Luthans, 2007	276	W	Stand-alone	Voluntary	Multiple	6 hr	Both	Many	Experimental	.29 (.13)
Cornett-DeVito & McGlone, 2000	30	W	Stand-alone	Mandatory	Inclusive	8 hr	Both	Many	Observational	.42 (.19)
D'Andrea et al., 1991										
Sample a	34	E	Stand-alone	Mandatory	One	15 weeks	Both	Many	Quasi-experimental	1.13 (.36)
Sample b	29	E	Stand-alone	Mandatory	One	6 weeks	Both	Many	Quasi-experimental	.79 (.39)
Sample c	27	E	Stand-alone	Mandatory	One	3 weeks	Both	Many	Observational	.75 (.21)
Darnell & Cook, 2009	98	E	Stand-alone	Voluntary	One	2 hr	Awareness	One	Quasi-experimental	.09 (.20)
Darst, 1988	20	W	Stand-alone	Voluntary	One	10 hr	Both	Many	Quasi-experimental	.38 (.31)
Davidson, 1991	40	E	Stand-alone	Voluntary	One	15 weeks	Both	Many	Quasi-experimental	.27 (.32)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A (*continued*)

Author, year	*N*	Training context: Setting	Training context: Training approach	Training context: Training attendance	Training design: Training focus	Training design: Training duration	Training design: Training type	Training design: Instruction	Methodological moderators study rigor	*g* (*SE*)
de Lemus, Navarro, Megias, Velasquez, & Ryan, 2013										
Study 1	28	E	Stand-alone	Voluntary	One	20 hr	Both	Many	Observational	.67 (.25)
Study 2	83	E	Stand-alone	Voluntary	One	20 hr	Both	Many	Quasi-experimental	.38 (.32)
Study 3	67	E	Stand-alone	Voluntary	One	21 hr	Both	Many	Quasi-experimental	.21 (.18)
De Meuse, Hostager, & O'Neill, 2007	35	W	Stand-alone	Unknown	Inclusive	1.5 days	Both	Many	Observational	.58 (.19)
Delgado et al., 2013	98	W	Stand-alone	Voluntary	One	1 hr	Awareness	One	Observational	.13 (.12)
Devine et al., 2012	91	E	Stand-alone	Voluntary	One	12 weeks	Awareness	One	Experimental	.11 (.21)
Diaz-Lazaro & Cohen, 2001	15	E	Stand-alone	Voluntary	Multiple	Unknown	Both	Many	Observational	.79 (.30)
Dickson, Argus-Calvo, & Tafoya, 2010	60	E	Stand-alone	Voluntary	One	15 weeks	Awareness	Many	Quasi-experimental	1.34 (.30)
Dogra, 2001	140	E	Integrated	Mandatory	Multiple	6 hr	Both	Many	Observational	−.01 (.08)
Dorfman, Murty, Ingram, & Li, 2007	167	E	Stand-alone	Mandatory	One	Course	Awareness	Many	Observational	.37 (.08)
Drwecki, 2011	60	E	Stand-alone	Voluntary	One	Unknown	Behaviors	One	Experimental	.27 (.26)
Dyson, 2003	115	E	Stand-alone	Voluntary	Inclusive	90 min	Both	Many	Experimental	.30 (.19)
Earley, 1987										
Training 1	80	W	Integrated	Mandatory	One	3 days	Both	Many	Experimental	2.07 (.39)
Training 2	80	W	Integrated	Mandatory	One	3 days	Awareness	One	Experimental	1.19 (.34)
Training 3	80	W	Integrated	Mandatory	One	3 days	Both	One	Experimental	1.14 (.34)
Edwards, 1997	48	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	.43 (.29)
Edwards, 1999	160	E	Stand-alone	Voluntary	One	Semester	Awareness	Many	Quasi-experimental	.41 (.16)
Ellison, 2002	52	E	Stand-alone	Mandatory	Inclusive	17.5 hr	Both	Many	Quasi-experimental	.49 (.28)
Finken, 2002	280	E	Stand-alone	Voluntary	One	Semester	Awareness	Many	Quasi-experimental	.40 (.12)
Fischer, 2011	49	E	Stand-alone	Voluntary	One	6 weeks	Awareness	Many	Observational	.01 (.14)
Flamini, 2005	71	E	Stand-alone	Voluntary	One	Semester	Awareness	Many	Quasi-experimental	.23 (.24)
Fluck, 2003	62	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	1.80 (.30)
Fradkin, 1980	30	E	Stand-alone	Voluntary	One	5 weeks	Both	Many	Experimental	−.14 (.36)
Gannon & Poon, 1997	105	E	Stand-alone	Voluntary	One	1 day	Both	Many	Observational	.30 (.10)
Gany & Thiel de Bocanegra, 1996	80	W	Stand-alone	Voluntary	Multiple	7.5 hr	Both	Many	Observational	.36 (.11)
Garcia, 1996	74	E	Stand-alone	Voluntary	One	12 hr	Both	Many	Experimental	.42 (.25)
Gerla, 1999	96	E	Stand-alone	Voluntary	One	75 min	Awareness	One	Observational	−.05 (.14)
Gharib & Phillips, 2012	60	E	Stand-alone	Voluntary	One	6 weeks	Awareness	One	Observational	.89 (.15)
Glaser, 2006	123	E	Stand-alone	Voluntary	Inclusive	20 min	Awareness	One	Experimental	.54 (.18)
Goldberg, 1982	131	E	Stand-alone	Voluntary	One	1 day	Awareness	One	Experimental	−.23 (.24)
Goldstein & Smith, 1999	81	W	Stand-alone	Unknown	Inclusive	1 week	Awareness	Many	Quasi-experimental	.57 (.23)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A (*continued*)

Author, year	*N*	Training context: Setting	Training context: Training approach	Training context: Training attendance	Training design: Training focus	Training design: Training duration	Training design: Training type	Training design: Instruction	Methodological moderators study rigor	*g* (*SE*)
Gorton, 1981	131	E	Stand-alone	Voluntary	One	1 day	Awareness	One	Quasi-experimental	.78 (.21)
Govern, 1997	140	E	Stand-alone	Mandatory	One	30 min	Awareness	One	Quasi-experimental	−.17 (.17)
Grant, 2003	39	E	Stand-alone	Voluntary	One	Semester	Both	Many	Experimental	.35 (.33)
Griswold et al., 2006	95	E	Stand-alone	Voluntary	One	1 year	Awareness	Many	Observational	1.30 (.14)
Gursimsek, 2010	478	E	Stand-alone	Unknown	One	Course	Awareness	Many	Quasi-experimental	−.41 (.10)
Guth, Lopez, Rojas, Clements, & Tyler, 2004	50	E	Stand-alone	Voluntary	One	7 weeks	Awareness	Many	Quasi-experimental	−.26 (.35)
Guy-Walls, 2007	150	E	Integrated	Mandatory	One	4 years	Both	Many	Quasi-experimental	.86 (.21)
Hammer & Martin, 1992										
Training 1	211	W	Stand-alone	Voluntary	One	15 hr	Awareness	Many	Quasi-experimental	.27 (.29)
Training 2	111	W	Stand-alone	Voluntary	One	30 hr	Both	Many	Quasi-experimental	1.11 (.31)
Handler, 1999	138	E	Stand-alone	Voluntary	One	90 min	Awareness	One	Observational	3.25 (.21)
Hanover & Cellar, 1998	99	W	Stand-alone	Mandatory	Multiple	30 min	Both	Many	Quasi-experimental	.53 (.20)
Harrison, 1992	65	W	Stand-alone	Voluntary	Inclusive	Unknown	Both	Many	Experimental	1.11 (.44)
Hauenstein et al., 2010	46	W	Stand-alone	Voluntary	One	15 weeks	Awareness	Many	Observational	.35 (.14)
Hayes et al., 2004										
Training 1	85	W	Stand-alone	Voluntary	Inclusive	1 day	Awareness	Many	Experimental	.24 (.26)
Training 2	85	W	Stand-alone	Voluntary	Multiple	1 day	Awareness	Many	Experimental	.09 (.25)
Henderson-King & Kaleta, 2000	385	E	Stand-alone	Mandatory	Multiple	Semester	Awareness	Many	Quasi-experimental	.00 (.11)
Hernandez & Gonzalez, 2008	86	E	Stand-alone	Voluntary	One	32 hr	Awareness	Many	Quasi-experimental	−1.77 (.28)
Hill & Augoustinos, 2001	62	W	Stand-alone	Voluntary	One	3 days	Both	Many	Observational	.46 (.20)
Hilliard, 2011	25	E	Stand-alone	Voluntary	One	Semester	Awareness	One	Observational	.31 (.20)
Hillman & Martin, 2002	68	E	Stand-alone	Voluntary	One	1 hr	Awareness	One	Experimental	.15 (.27)
Ho, Yao, Lee, Beach, & Green, 2008	57	E	Stand-alone	Mandatory	One	2 hr	Both	Many	Experimental	.73 (.33)
Hodson, Choma, & Costello, 2009	128	E	Integrated	Voluntary	One	40 min	Awareness	One	Experimental	.42 (.24)
Hoff, 2005	20	E	Stand-alone	Voluntary	One	Semester	Both	Many	Experimental	.16 (43)
Hogan & Mallott, 2005	250	E	Unknown	Mandatory	One	12 weeks	Unknown	Many	Quasi-experimental	.33 (.22)
Holladay & Quinones, 2008	165	W	Integrated	Voluntary	Inclusive	Unknown	Both	One	Observational	.70 (.12)
Hood, Muller, & Seitz, 2001	192	E	Stand-alone	Mandatory	Multiple	16 weeks	Both	One	Observational	.31 (.08)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(*Appendices continue*)

Appendix A (*continued*)

Author, year	*N*	Training context: Setting	Training context: Training approach	Training context: Training attendance	Training design: Training focus	Training design: Training duration	Training design: Training type	Training design: Instruction	Methodological moderators study rigor	*g* (*SE*)
Hostager & De Meuse, 2008	177	E	Stand-alone	Voluntary	Multiple	16 weeks	Awareness	Many	Observational	.17 (.18)
Howe, 2001	41	E	Stand-alone	Voluntary	One	6 weeks	Awareness	Many	Observational	.35 (.17)
Hoyt, 1987	26	W	Stand-alone	Voluntary	One	45 min	Behaviors	One	Experimental	.55 (.39)
Hughes & Hood, 2007	218	E	Integrated	Mandatory	One	16 weeks	Both	Many	Observational	1.14 (.09)
Hughes-White, 1991	92	E	Stand-alone	Voluntary	One	Semester	Both	Many	Experimental	.42 (.21)
Hurtado, Mayhew, & Engberg, 2012	236	E	Stand-alone	Voluntary	Multiple	16 weeks	Both	Many	Quasi-experimental	.84 (.13)
Hussey & Bisconti, 2010										
Training 1	82	E	Stand-alone	Voluntary	One	2.25 hr	Awareness	Many	Observational	.29 (.03)
Training 2	82	E	Stand-alone	Voluntary	One	2 hr	Awareness	Many	Observational	.30 (.03)
Hussey, Fleck, & Warner, 2010	63	E	Stand-alone	Voluntary	Inclusive	Semester	Awareness	Many	Quasi-experimental	−.29 (.25)
Hylton, 2006	50	E	Stand-alone	Voluntary	One	15 weeks	Both	Many	Quasi-experimental	.27 (.03)
Ibrahim & Herr, 1976	50	E	Stand-alone	Voluntary	One	5 hr	Behavior	One	Experimental	.79 (.39)
Israel, 1998										
Training 1	161	E	Stand-alone	Voluntary	One	2.5 hr	Awareness	Many	Experimental	1.07 (.26)
Training 2	161	E	Stand-alone	Voluntary	One	2.5 hr	Awareness	Many	Experimental	−.12 (.22)
Training 3	161	E	Stand-alone	Voluntary	One	2.5 hr	Awareness	Many	Experimental	.17 (.24)
Jefferson, 2001	261	E	Stand-alone	Mandatory	One	4 hr	Both	Many	Observational	−.05 (.06)
Jeffreys & Dogan, 2012	36	E	Stand-alone	Voluntary	Multiple	2 years	Both	Many	Observational	.61 (.18)
Johnson, Antle, & Barbee, 2009	462	W	Stand-alone	Voluntary	One	2.5 days	Awareness	Many	Observational	−.34 (.05)
Jones, 1991	141	W	Stand-alone	Voluntary	One	2 days	Both	Many	Experimental	.26 (.23)
Jones, 2008	20	W	Stand-alone	Voluntary	One	13 weeks	Both	Many	Observational	.06 (.22)
Juarez et al., 2006	11	W	Stand-alone	Unknown	Multiple	36 hr	Both	Many	Observational	.87 (.37)
Kamfer & Venter, 1994	37	E	Stand-alone	Voluntary	One	Unknown	Awareness	One	Quasi-experimental	−.84 (.34)
Katz, 1977	24	E	Stand-alone	Mandatory	One	24 hr	Both	Many	Quasi-experimental	.37 (.34)
Katz & Ivey, 1977										
Sample 1	24	E	Integrated	Mandatory	One	26 hr	Awareness	Many	Observational	1.31 (.38)
Sample 2	24	E	Integrated	Mandatory	One	26 hr	Awareness	Many	Observational	1.27 (.37)
Keillor, 1999	81	E	Stand-alone	Voluntary	One	6 weeks	Both	Many	Quasi-experimental	.21 (.22)
Keim, Warring, & Rau, 2001	63	E	Stand-alone	Voluntary	Multiple	15 weeks	Both	Many	Observational	.86 (.18)
Kennedy, 1995	253	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	.18 (.14)
Kernahan & Davis, 2007	39	E	Integrated	Mandatory	One	Semester	Awareness	Many	Quasi-experimental	1.40 (.36)
Kernahan & Davis, 2010	57	E	Stand-alone	Voluntary	One	Semester	Both	Many	Observational	.27 (.28)
Kilmnick, 2006	140	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	−.47 (.17)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(*Appendices continue*)

Appendix A (*continued*)

Author, year	*N*	Training context			Training design				Methodological moderators study rigor	*g* (*SE*)
		Setting	Training approach	Training attendance	Training focus	Training duration	Training type	Instruction		
Kitchens-Stephens, 2005	16	E	Stand-alone	Voluntary	One	Unknown	Both	Many	Quasi-experimental	8.27 (1.54)
Klak & Martin, 2003	63	E	Integrated	Unknown	One	Semester	Awareness	Many	Observational	.14 (.13)
Kohl, 2005	12	E	Stand-alone	Voluntary	One	2 Semesters	Both	Many	Quasi-experimental	.82 (.32)
Kracht, 1998	141	W	Stand-alone	Voluntary	Inclusive	1 day	Awareness	Many	Observational	.22 (.08)
Lal, 2010	34	W	Stand-alone	Voluntary	One	1.5 hr	Both	Many	Quasi-experimental	1.24 (.23)
Landis, Brislin, & Hulgus, 1985	45	E	Stand-alone	Voluntary	One	30 min	Awareness	One	Experimental	.45 (.45)
Landis, Day, McGrew, Thomas, & Miller, 1976	170	W	Stand-alone	Voluntary	One	Unknown	Awareness	One	Observational	.05 (.07)
Law, 1998	66	W	Stand-alone	Voluntary	One	1.5 hr	Both	Many	Quasi-experimental	.56 (.25)
Lee et al., 2006	7	W	Stand-alone	Voluntary	One	90 min	Awareness	One	Observational	1.06 (.43)
Liberman, Block, & Koch, 2011										
Study 1	124	E	Stand-alone	Voluntary	Unknown	Unknown	Unknown	Many	Experimental	.51 (.20)
Study 2	100	E	Stand-alone	Voluntary	Unknown	Unknown	Unknown	Many	Experimental	.27 (.21)
Lichtenstein, Lindstrom, & Povenmire-Kirk, 2008	53	W	Stand-alone	Voluntary	One	1.5 hr	Awareness	One	Quasi-experimental	.35 (.31)
Livosky, Pettijohn, & Capo, 2011	27	E	Stand-alone	Voluntary	One	10 weeks	Awareness	Many	Observational	.53 (.20)
LoboPrabhu, King, Albucher, & Liberzon, 2000	24	E	Stand-alone	Voluntary	One	4 hr	Both	Many	Observational	.74 (.23)
Lopez-Humphreys, 2011	115	E	Stand-alone	Voluntary	One	14 weeks	Both	Many	Quasi-experimental	.14 (.20)
Luger, 2011	24	E	Stand-alone	Voluntary	One	unknown	Both	Many	Observational	.14 (.20)
Madera, Neal, & Dawson, 2011	96	E	Stand-alone	Voluntary	One	1 day	Awareness	One	Observational	.20 (.10)
Majumdar et al., 2004	69	W	Stand-alone	Voluntary	One	36 hr	Awareness	One	Experimental	.41 (.12)
Majumdar, Keystone, & Cuttress, 1999	48	E	Stand-alone	Voluntary	Inclusive	15 hr	Both	Many	Quasi-experimental	.52 (.33)
Mak & Buckingham, 2007	142	E	Integrated	Voluntary	One	6 weeks	Both	Many	Quasi-experimental	.47 (.29)
Manese, Wu, & Nepomuceno, 2001	24	E	Integrated	Mandatory	Multiple	1 year	Both	Many	Observational	.33 (.21)
Manis, 2008	99	E	Stand-alone	Voluntary	One	3 hr	Both	Many	Quasi-experimental	.09 (.20)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A (*continued*)

Author, year	*N*	Training context: Setting	Training context: Training approach	Training context: Training attendance	Training design: Training focus	Training design: Training duration	Training design: Training type	Training design: Instruction	Methodological moderators: study rigor	*g* (*SE*)
Mapp, McFarland, & Newell, 2007	23	E	Stand-alone	Voluntary	One	2 weeks	Awareness	Many	Observational	.14 (.20)
Martin, 2006	138	E	Stand-alone	Voluntary	Inclusive	Semester	Both	Many	Quasi-experimental	.32 (.15)
Martinez, 1995	29	E	Stand-alone	Voluntary	One	2 days	Both	Many	Quasi-experimental	.23 (.38)
Mausehund, Timm, & King, 1995	67	E	Stand-alone	Mandatory	Inclusive	50 min	Awareness	Many	Quasi-experimental	.03 (.24)
May, 2010	91	E	Stand-alone	Voluntary	One	30 min	Awareness	One	Quasi-experimental	.36 (.21)
McCleskey, 1991	87	E	Stand-alone	Voluntary	One	2 hr	Both	Many	Quasi-experimental	.54 (.24)
McCool, Du Toit, Petty, & McCauley, 2006	84	E	unknown	Voluntary	One	Unknown	Awareness	Many	Observational	−.37 (.05)
Medina-Walpole, Mooney, Lyness, Lambert, & Lurie, 2012	101	E	Stand-alone	Mandatory	One	10 weeks	Both	Many	Observational	.20 (.10)
Middleton, 2002	79	E	Integrated	Voluntary	Multiple	15 weeks	Both	Many	Observational	.42 (.12)
Mio, 1989	27	E	Stand-alone	Mandatory	One	Semester	Both	One	Quasi-experimental	.79 (.40)
Moffat & Tung, 2004	49	E	Stand-alone	Voluntary	Inclusive	2 days	Awareness	Many	Observational	.44 (.15)
Morin, 1974	18	E	Stand-alone	Voluntary	One	40 hr	Unknown	Many	Observational	.29 (.23)
Moss-Racusin et al., 2013	126	W	Stand-alone	Voluntary	One	Unknown	Unknown	One	Observational	.47 (.19)
Moyer & Nath, 1998	84	E	Stand-alone	Voluntary	One	30 min	Awareness	One	Experimental	.73 (.24)
Murphy et al., 2006	12	E	Integrated	Mandatory	Multiple	16 weeks	Awareness	Many	Observational	1.28 (.38)
Mysore, 2004	72	E	Stand-alone	Voluntary	One	4 weeks	Awareness	Many	Quasi-experimental	.52 (.25)
Neville & Furlong, 1994	75	E	Stand-alone	Voluntary	One	2 hr	Both	Many	Experimental	.35 (.23)
Neville et al., 1996	38	E	Stand-alone	Voluntary	One	15 weeks	Awareness	Many	Observational	.61 (.20)
Nguyen, Tran, Derous, Lopez, & Siu, 2012	368	E	Stand-alone	Voluntary	One	Unknown	Both	Many	Experimental	.27 (.10)
Nokes, Nickitas, Keida, & Neville, 2005	14	E	Stand-alone	Voluntary	Inclusive	15 hr	Awareness	Many	Observational	−.16 (.32)
Paluck, 2009	40	W	Stand-alone	Voluntary	Inclusive	1 year	Awareness	One	Experimental	.23 (.31)
Parker, Moore, & Neimeyer, 1998	58	E	Stand-alone	Mandatory	One	15 weeks	Both	Many	Experimental	.39 (.28)
Pedersen, 2010	45	E	Integrated	Voluntary	Inclusive	Semester	Awareness	Many	Quasi-experimental	.61 (.28)
Pedersen & Barlow, 2008	123	E	Stand-alone	Voluntary	One	6 weeks	Awareness	Many	Observational	−.49 (.14)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(*Appendices continue*)

Appendix A (*continued*)

Author, year	*N*	Training context			Training design				Methodological moderators study rigor	*g* (*SE*)
		Setting	Training approach	Training attendance	Training focus	Training duration	Training type	Instruction		
Perry, Kulik, & Schmidtke, 1998	36	E	Stand-alone	Voluntary	One	20 min	Awareness	One	Observational	.07 (.34)
Pettijohn & Walzer, 2008	99	E	Stand-alone	Unknown	Multiple	10 weeks	Awareness	Many	Observational	.53 (.20)
Pilkington, 1993	59	W	Stand-alone	Voluntary	One	12 weeks	Both	Many	Quasi-experimental	.34 (.26)
Preusser, Bartels, & Nordstrom, 2011	70	W	Stand-alone	Voluntary	One	2 hr	Awareness	One	Quasi-experimental	.39 (.24)
Probst, 2003	94	E	Stand-alone	Voluntary	Multiple	17 weeks	Awareness	Many	Quasi-experimental	.30 (.21)
Randolph, Landis, & Tzeng, 1977	35	E	Stand-alone	Voluntary	One	Unknown	Both	One	Experimental	.74 (.41)
Rehg, Gundalch, & Grigorian, 2012	110	W	Stand-alone	Voluntary	One	9 days	Awareness	One	Observational	.55 (.17)
Reinhardt, 1994	320	E	Stand-alone	Voluntary	One	1 hr	Awareness	One	Quasi-experimental	.26 (.11)
Remer, 2008	278	E	Stand-alone	Voluntary	Inclusive	1 semester	Awareness	Many	Quasi-experimental	.18 (.13)
Remmert, 1993	103	W	Stand-alone	Voluntary	One	15 weeks	Both	Many	Quasi-experimental	.20 (.20)
Reynolds, 2010	18	W	Stand-alone	Voluntary	One	4.5 hr	Both	Many	Observational	1.40 (.48)
Rhyne, 1973	72	W	Integrated	Unknown	One	8 weeks	Awareness	One	Experimental	.10 (.29)
Riggs, Rosenthal, & Smith-Bonahue, 2011	67	E	Stand-alone	Voluntary	One	3 hr	Awareness	Many	Experimental	1.23 (.23)
Roberson et al., 2002										
Condition 1	98	E	Stand-alone	Voluntary	One	Unknown	Awareness	One	Observational	.00 (.15)
Condition 2	98	E	Stand-alone	Voluntary	One	Unknown	Behavior	One	Observational	.56 (.16)
Robinson & Bradley, 1997	44	E	Stand-alone	Voluntary	One	3 weeks	Awareness	Many	Quasi-experimental	.90 (.28)
Rogers-Sirin & Sirin, 2009	95	E	Stand-alone	Voluntary	One	1 day	Both	Many	Observational	.29 (.10)
Rouh, 2001	130	W	Stand-alone	Mandatory	One	2 hr	Awareness	Many	Quasi-experimental	1.20 (.19)
Rousey, 2010	118	E	Stand-alone	Voluntary	One	45 min	Awareness	One	Quasi-experimental	.05 (.18)
Rowell & Benshoff, 2008	199	E	Integrated	Voluntary	Inclusive	Unknown	Awareness	One	Observational	.10 (.11)
Rudman et al., 2001										
Study 1	107	E	Stand-alone	Voluntary	One	14 weeks	Both	Many	Quasi-experimental	.95 (.36)
Study 2	107	E	Stand-alone	Voluntary	One	14 weeks	Both	Many	Quasi-experimental	.55 (.25)
Rudolph, 1989	52	W	Stand-alone	Voluntary	One	3 days	Both	Many	Quasi-experimental	.49 (.28)
Sakurai, McCall-Wolf, & Kashima, 2010	98	E	Stand-alone	Voluntary	One	1 hr	Unknown	One	Quasi-experimental	.46 (.20)
Sanchez & Medkik, 2004	125	W	Stand-alone	Mandatory	One	1 day	Awareness	Many	Quasi-experimental	.16 (.18)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A (*continued*)

Author, year	*N*	Training context: Setting	Training context: Training approach	Training context: Training attendance	Training design: Training focus	Training design: Training duration	Training design: Training type	Training design: Instruction	Methodological moderators study rigor	*g* (*SE*)
Sanchez-Burks, Lee, Nisbett, & Ybarra, 2007	74	E	Stand-alone	Unknown	One	Unknown	Both	Many	Quasi-experimental	.70 (.24)
Sanner, Baldwin, Cannella, Charles, & Parker, 2010	47	W	Stand-alone	Voluntary	Inclusive	3 hr	Awareness	One	Observational	.50 (.15)
Scher, 2008	69	E	Stand-alone	Voluntary	One	2.25 hr	Both	Many	Observational	.80 (.22)
Seguin, 2002	251	W	Stand-alone	Voluntary	One	Unknown	Both	Many	Observational	.08 (.06)
Shergill, 1997	36	E	Stand-alone	Voluntary	One	3 weeks	Behavior	Many	Experimental	.21 (.32)
Shields, Zawadzki, & Johnson, 2011	118	E	Stand-alone	Voluntary	One	Unknown	Awareness	One	Experimental	.64 (.19)
Soble, Spanierman, & Liao, 2011	138	E	Stand-alone	Voluntary	One	20 min	Awareness	One	Experimental	.13 (.18)
St. Clair & McKenry, 1999	200	E	Integrated	Voluntary	Inclusive	3 weeks	Both	Many	Quasi-experimental	.30 (.13)
Stebbins, 2005	106	E	Stand-alone	Voluntary	Inclusive	8 hr	Awareness	Many	Experimental	.29 (.19)
Steed, 2010	13	W	Stand-alone	Voluntary	One	6 hr	Awareness	Many	Observational	−.07 (.27)
Steinfeldt & Wong, 2010	43	E	Stand-alone	Voluntary	One	45 min	Awareness	One	Quasi-experimental	.56 (.31)
Stella, Forlin, & Lan, 2007	151	E	Stand-alone	Voluntary	One	10 weeks	Both	Many	Observational	.24 (.08)
Stewart et al., 2003	34	E	Stand-alone	Voluntary	One	8 hr	Both	Many	Experimental	.64 (.37)
Stewart, Latu, Kawakami, & Myers, 2010	72	E	Stand-alone	Voluntary	One	Unknown	Awareness	One	Quasi-experimental	.53 (.24)
Stowe, 2002	27	E	Stand-alone	Voluntary	One	10 hr	Both	Many	Experimental	.16 (.37)
Tang, Fantone, Bozynski, & Adams, 2002	148	E	Integrated	Mandatory	One	Unknown	Both	Many	Observational	.12 (.08)
Tang, Hernandez, & Adams, 2004	12	E	Stand-alone	Mandatory	One	3 hr	Both	Many	Observational	.36 (.28)
Taras et al., 2013	1299	E	Stand-alone	Voluntary	Inclusive	Unknown	Awareness	Many	Quasi-experimental	.04 (.02)
Taton, 2008	204	E	Stand-alone	Voluntary	One	Unknown	Awareness	Many	Experimental	.19 (.25)
Taylor-Ritzler et al., 2008	287	W	Stand-alone	Voluntary	One	7 hr	Awareness	Many	Observational	.16 (.06)
Theis-Cole, 1995	153	E	Stand-alone	Voluntary	One	2.5 hr	Both	One	Quasi-experimental	.14 (.16)
Thomas & Cohn, 2006)	47	W	Stand-alone	Voluntary	Multiple	9 days	Both	Many	Observational	.61 (.18)
Thomas, 2008	88	W	Stand-alone	Voluntary	One	Unknown	Unknown	One	Quasi-experimental	.29 (.22)
Timm, 1997	74	E	Stand-alone	Voluntary	One	2 hr	Both	Many	Quasi-experimental	.30 (.23)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A (*continued*)

Author, year	*N*	Training context: Setting	Training context: Training approach	Training context: Training attendance	Training design: Training focus	Training design: Training duration	Training design: Training type	Training design: Instruction	Methodological moderators study rigor	*g* (*SE*)
Todd, Bodenhausen, Richeson, & Galinsky, 2011	56	E	Stand-alone	Voluntary	One	1 day 6 hr	Awareness	One	Experimental	.59 (.28)
Tolleson Knee, 1999	208	E	Stand-alone	Voluntary	Multiple	15 weeks	Both	One	Quasi-experimental	.11 (.17)
Tran et al., 1994	55	E	Stand-alone	Voluntary	One	15 weeks	Both	Many	Observational	.13 (.13)
Turner, 1986	169	W	Stand-alone	Voluntary	One	Unknown	Both	Many	Quasi-experimental	.20 (.35)
Van Soest, 1996	222	E	Stand-alone	Voluntary	Multiple	Semester	Awareness	Many	Quasi-experimental/ Observataional	−.23 (.14)
VonDras & Lor-Vang, 2004										
Sample a	30	E	Integrated	Mandatory	One	Unknown	Awareness	One	Observational	.24 (.18)
Sample b	41	E	Integrated	Mandatory	One	Unknown	Awareness	One	Observational	.28 (.16)
Sample c	64	E	Integrated	Mandatory	One	Unknown	Awareness	One	Observational	.21 (.12)
Wade & Bernstein, 1991	8	W	Stand-alone	Voluntary	One	4 hr	Both	Many	Quasi-experimental	.24 (.23)
Waight & Madera, 2011	186	W	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Observational	−.02 (.23)
Wang, 1998	62	E	Stand-alone	Voluntary	Inclusive	10 days	Awareness	One	Quasi-experimental	.85 (.26)
Warring, Keim, & Rau, 1998	136	E	Stand-alone	Mandatory	Inclusive	15 weeks	Both	Many	Observational	.82 (.13)
Weathersby, 2004	310	E	Stand-alone	Mandatory	One	Semester	Awareness	Many	Quasi-experimental	.57 (.12)
Wideman, 2005	41	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	.19 (.31)
Wiggins, Follo, & Eberly, 2007	47	E	Stand-alone	Voluntary	One	60 hr	Both	Many	Quasi-experimental	.61 (.30)
Williams, 2005	47	W	Integrated	Voluntary	Inclusive	1 month	Both	Many	Quasi-experimental	.35 (.30)
Worchel & Mitchell, 1972	127	W	Stand-alone	Unknown	One	3 hr	Awareness	One	Experimental	.38 (.20)
Wortman, 2002	188	E	Stand-alone	Voluntary	One	Semester	Both	Many	Quasi-experimental	.12 (.15)
Zarubin, 2008	242	E	Stand-alone	Mandatory	Multiple	1 hr	Behavior	One	Observational	.22 (.10)

Note. E = educational; W = workplace.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B

Effect Size Estimate per Outcome Variable

Author, year	Outcome variable	Training effects	*g* (*SE*)
Abernethy, 2005	Attitude	Short	.87 (.29)
	Behavior	Short	.72 (.28)
	Cognition	Short	.93 (.30)
Aldridge, 2013	Attitude	Short	.65 (.27)
	Behavior	Short	.69 (.27)
Alonso, 2005			
Sample a	Attitude	Short	.96 (.30)

(*Appendices continue*)

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Sample a	Behavior	Short	.56 (.27)
Sample a	Cognition	Short	.95 (.30)
Sample b	Attitude	Short	.65 (.21)
Sample b	Behavior	Short	.64 (.21)
Sample b	Cognition	Short	.86 (.22)
Altshuler et al., 2003			
Training 1	Attitude	Long	.00 (.29)
Training 2	Attitude	Short	−.18 (.33)
Amatea et al., 2012	Attitude	Short	.61 (.09)
	Reaction	Short	1.14 (.11)
Anderson, 1981	Attitude	Short	.73 (.27)
Arizaga, 1999	Attitude	Short	.00 (.26)
	Behavior	Short	.74 (.28)
Armour et al., 2004	Behavior	Long	.88 (.34)
Baba & Hebert, 2004	Attitude	Short	−.18 (.08)
Bailey et al., 2001	Attitude	Short	.47 (.27)
Baker, 2009	Attitude	Short	.00 (.15)
Ballou, 1996	Attitude	Short	.23 (.11)
Barker & Hartel, 2004	Behavior	Short	.77 (.26)
Bauer et al., 2013	Attitude	Short	−.46 (.14)
Bhawuk, 1998	Attitude	Short	.77 (.29)
Blakely et al., 1998	Cognition	Long	.26 (.16)
Bloch, 2012	Attitude	Long	−.44 (.22)
	Attitude	Short	.81 (.24)
	Behavior	Long	−.44 (.22)
	Behavior	Short	.81 (.24)
	Cognition	Long	−.44 (.22)
	Cognition	Short	.81 (.24)
Boswell, 2012	Attitude	Short	.53 (.16)
	Cognition	Short	.53 (.16)
	Behavior	Short	.53 (.16)
Brathwaite & Majumdar, 2006	Cognition	Long	.27 (.12)
	Cognition	Short	.27 (.12)
Brown et al., 1996	Attitude	Short	.64 (.39)
	Behavior	Short	.77 (.37)
	Cognition	Short	.28 (.35)
Brown, 2004	Cognition	Short	−.13 (.20)
Burris, 1992	Cognition	Short	1.84 (.29)
Bush & Ingram, 2001	Attitude	Short	.37 (.11)
Butler et al., 2012	Attitude	Short	.32 (.23)
	Reaction	Short	.48 (.23)
Byington et al., 1997	Attitude	Short	.80 (.16)
	Behavior	Short	.23 (.15)
Caffrey et al., 2005	Attitude	Short	.22 (.42)
	Behavior	Short	1.48 (.48)
	Cognition	Short	1.37 (.45)
	Overall	Short	1.92 (.48)
Cap, 1995	Attitude	Short	.23 (.27)
	Cognition	Short	1.16 (.29)
Carpenter, 1997	Behavior	Short	1.24 (.28)
Carrell, 1997			
Sample a	Attitude	Short	−.14 (.25)
Sample a	Behavior	Short	−.16 (.25)
Sample b	Attitude	Short	.44 (.22)
Sample b	Behavior	Short	.56 (.23)
Sample c	Attitude	Short	.15 (.34)
Sample c	Behavior	Short	.35 (.35)
Sample d	Attitude	Short	.56 (.26)
Sample d	Behavior	Short	1.20 (.28)
Carter et al., 2006	Attitude	Long	.01 (.07)
Cascio & Bass, 1976	Attitude	Short	.04 (.02)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(*Appendices continue*)

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Case, 2007	Attitude	Short	.03 (.08)
Case & Stewart, 2010	Attitude	Short	.42 (.22)
Case & Stewart, 2013	Attitude	Short	.41 (.09)
Castillo et al., 2007	Attitude	Short	.73 (.22)
	Behavior	Short	.05 (.22)
	Cognition	Short	.09 (.22)
Cates, 2006	Attitude	Short	1.01 (.31)
	Cognition	Short	.92 (.30)
Celik, Abma, Klinge, & Widdershoven, 2012	Attitude	Short	.64 (.33)
	Cognition	Short	1.88 (.72)
Chang, 2002	Attitude	Short	.17 (.16)
Chrobot-Mason, 2004	Attitude	Short	.13 (.28)
	Behavior	Short	.72 (.28)
Clark, 1998	Attitude	Short	.18 (.16)
Clevenger, 2011	Cognition	Short	.37 (.12)
Clinton, 1983	Cognition	Short	.12 (.31)
	Attitude	Short	.60 (.31)
Cole et al., 2011	Attitude	Short	.24 (.16)
	Behavior	Short	.08 (.15)
	Cognition	Short	.02 (.15)
Colvin-Burque et al., 2007	Attitude	Short	.47 (.10)
Combs & Luthans, 2007	Attitude	Short	.28 (.11)
	Behavior	Long	.28 (.20)
	Behavior	Short	.30 (.07)
Cornett-DeVito & McGlone, 2000	Behavior	Short	.42 (.19)
D'Andrea et al., 1991			
Sample a	Attitude	Short	1.31 (.37)
Sample a	Behavior	Short	1.25 (.37)
Sample a	Cognition	Short	.81 (.35)
Sample b	Attitude	Short	.82 (.39)
Sample b	Behavior	Short	1.00 (.40)
Sample b	Cognition	Short	.56 (.38)
Sample c	Attitude	Short	.86 (.22)
Sample c	Behavior	Short	.86 (.22)
Sample c	Cognition	Short	.53 (.20)
Darnell & Cook, 2009	Attitude	Short	.09 (.02)
Darst, 1988	Cognition	Short	.38 (.31)
Davidson, 1991	Cognition	Short	.26 (.32)
	Attitude	Short	.29 (.32)
de Lemus et al., 2013			
Study 1	Behavior	Short	.49 (.24)
Study 1	Attitude	Short	.86 (.27)
Study 2	Behavior	Short	.32 (.32)
Study 2	Attitude	Short	.44 (.33)
Study 3	Behavior	Short	.00 (.12)
Study 3	Attitude	Short	.42 (.25)
De Meuse et al., 2007	Attitude	Long	.34 (.18)
	Attitude	Short	.53 (.18)
	Behavior	Long	.46 (.18)
	Behavior	Short	1.00 (.20)
Delgado et al., 2013	Cognition	Long	.13 (.12)
Devine et al., 2012	Attitude	Long	.51 (.39)
	Cognition	Long	.06 (.21)
Diaz-Lazaro & Cohen, 2001	Attitude	Short	.16 (.21)
	Behavior	Short	.8 (.30)
	Cognition	Short	.8 (.30)
Dickson et al., 2010	Attitude	Short	.80 (.30)
	Behavior	Short	.69 (.28)
	Cognition	Short	1.57 (.31)
Dogra, 2001	Attitude	Short	1.77 (.32)
	Behavior	Short	.05 (.09)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Dorfman et al., 2007	Attitude	Short	−.06 (.08)
	Behavior	Short	−.13 (.08)
	Cognition	Short	.48 (.08)
Drwecki, 2011	Attitude	Short	.76 (.09)
Dyson, 2003	Behavior	Short	.27 (.26)
	Cognition	Short	.24 (.19)
Earley, 1987			
Training 1	Attitude	Long	.35 (.19)
Training 1	Behavior	Long	2.24 (.40)
Training 2	Attitude	Long	1.90 (.38)
Training 2	Behavior	Long	1.31 (.34)
Training 3	Attitude	Long	1.07 (.33)
Training 3	Behavior	Long	.91 (.33)
Edwards, 1997	Attitude	Short	1.38 (.35)
	Behavior	Short	.68 (.29)
	Cognition	Short	.53 (.29)
Edwards, 1999	Attitude	Short	.08 (.28)
Ellison, 2002	Attitude	Short	.41 (.16)
Finken, 2002	Attitude	Short	.40 (.12)
Fischer, 2011	Attitude	Short	.50 (.28)
	Cognition	Short	.28 (.14)
Flamini, 2005	Attitude	Short	−.27 (.14)
Fluck, 2003	Attitude	Short	.23 (.24)
Fradkin, 1980	Attitude	Long	1.80 (.30)
	Attitude	Short	−.20 (.36)
Gannon & Poon, 1997	Attitude	Short	−.09 (.36)
Gany & Thiel de Bocanegra, 1996	Attitude	Long	.34 (.11)
	Cognition	Long	.38 (.12)
Garcia, 1996	Attitude	Short	.37 (.25)
	Cognition	Short	.56 (.25)
	Behavior	Short	.34 (.25)
Gerla, 1999	Attitude	Long	−.02 (.14)
	Attitude	Short	−.08 (.14)
Gharib & Phillips, 2012	Cognition	Short	.90 (.15)
	Reaction	Short	.87 (.14)
Glaser, 2006	Attitude	Short	.54 (.18)
Goldberg, 1982	Attitude	Short	−.28 (.25)
	Attitude	Long	−.19 (.24)
Goldstein & Smith, 1999	Attitude	Short	.57 (.23)
Gorton, 1981	Attitude	Short	.78 (.21)
Govern, 1997	Attitude	Short	−.17 (.17)
Grant, 2003	Attitude	Short	.35 (.33)
Griswold et al., 2006	Cognition	Short	1.30 (.14)
Gursimsek, 2010	Attitude	Short	−.41 (.10)
Guth et al., 2004	Attitude	Short	−.42 (.35)
	Attitude	Long	−.10 (.34)
Guy-Walls, 2007	Attitude	Short	.86 (.21)
	Behavior	Short	.86 (.21)
	Cognition	Short	.86 (.21)
Hammer & Martin, 1992			
Training 1	Attitude	Short	.35 (.34)
	Cognition	Short	.19 (.23)
Training 2	Attitude	Short	1.38 (.37)
	Cognition	Short	.84 (.24)
Handler, 1999	Cognition	Short	3.25 (.21)
Hanover & Cellar, 1998	Attitude	Long	.60 (.20)
	Attitude	Short	.46 (.20)
Harrison, 1992	Attitude	Short	1.11 (.44)
Hauenstein et al., 2010	Behavior	Short	.35 (.14)
Hayes et al., 2004			
Training 1	Attitude	Long	.17 (.26)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Training 1	Attitude	Short	.30 (.26)
Training 2	Attitude	Long	−.15 (.25)
Training 2	Attitude	Short	.34 (.25)
Henderson-King & Kaleta, 2000	Attitude	Short	.00 (.11)
Hernandez & Gonzalez, 2008	Attitude	Short	−1.77 (.28)
Hill & Augoustinos, 2001	Cognition	Short	1.95 (.22)
	Attitude	Short	.59 (.14)
	Cognition	Long	1.41 (.25)
	Attitude	Long	.23 (.18)
	Cognition	Long	−.80 (.20)
	Attitude	Long	−.61 (.20)
Hilliard, 2011	Attitude	Short	.31 (.20)
Hillman & Martin, 2002	Attitude	Short	.15 (.27)
Ho et al., 2008	Behavior	Short	.73 (.33)
Hodson et al., 2009	Attitude	Long	.39 (.24)
	Attitude	Short	.46 (.24)
Hoff, 2005	Attitude	Short	.16 (.43)
Hogan & Mallott, 2005	Attitude	Long	.01 (.25)
	Attitude	Short	.65 (.20)
Holladay & Quinones, 2008	Attitude	Long	.35 (.16)
	Attitude	Short	.83 (.09)
	Behavior	Long	.92 (.17)
	Cognition	Short	.86 (.09)
	Reaction	Short	.53 (.08)
Hood et al., 2001	Attitude	Short	.31 (.08)
Hostager & De Meuse, 2008	Cognition	Short	.18 (.18)
Howe, 2001	Attitude	Short	.35 (.17)
Hoyt, 1987	Attitude	Short	.55 (.39)
Hughes & Hood, 2007	Behavior	Short	1.15 (.09)
Hughes-White, 1991	Behavior	Short	.42 (.21)
Hurtado et al., 2012	Cognition	Short	.22 (.12)
	Reaction	Short	1.46 (.13)
Hussey et al., 2010	Attitude	Short	−.29 (.25)
Hussey & Bisconti, 2010			
Training 1	Behavior	Short	.42 (.03)
Training 1	Attitude	Short	.22 (.03)
Training 1	Cognition	Short	.23 (.03)
Training 2	Behavior	Short	.36 (.03)
Training 2	Attitude	Short	.29 (.03)
Training 2	Cognition	Short	.27 (.03)
Hylton, 2006	Attitude	Short	.01 (.30)
Ibrahim & Herr, 1976	Attitude	Short	.72 (.39)
	Attitude	Long	.86 (.40)
Israel, 1998			
Training 1	Cognition	Short	2.44 (.29)
Training 1	Attitude	Short	−.29 (.22)
Training 2	Cognition	Short	.05 (.22)
Training 2	Attitude	Short	−.29 (.22)
Training 3	Cognition	Short	1.29 (.24)
Training 3	Attitude	Short	−.96 (.24)
Jefferson, 2001	Attitude	Short	−.05 (.06)
Jeffreys & Dogan, 2012	Attitude	Long	.45 (.17)
	Behavior	Long	.50 (.17)
	Cognition	Long	.87 (.19)
Johnson et al., 2009	Attitude	Short	−.34 (.05)
Jones, 1991	Attitude	Short	.26 (.23)
Jones, 2008	Attitude	Short	.38 (.22)
	Cognition	Short	−.26 (.22)
Juarez et al., 2006	Behavior	Short	.88 (.35)
Kamfer & Venter, 1994	Attitude	Short	−.84 (.34)
Katz, 1977	Attitude	Short	1.31 (.48)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Katz & Ivey, 1977			
Sample 1	Attitude	Short	1.31 (.38)
Sample 1	Attitude	Long	.00 (.27)
Sample 2	Attitude	Short	1.27 (.37)
Sample 2	Behavior	Long	−.19 (.27)
Keillor, 1999	Attitude	Short	.31 (.22)
	Behavior	Short	.11 (.22)
Keim et al., 2001	Attitude	Short	.80 (.18)
	Behavior	Short	.74 (.18)
	Cognition	Short	1.05 (.19)
Kennedy, 1995	Attitude	Short	.18 (.14)
Kernahan & Davis, 2007	Attitude	Short	1.4 (.36)
Kernahan & Davis, 2010	Attitude	Long	−.16 (.25)
	Attitude	Short	.69 (.30)
Kilmnick, 2006	Attitude	Short	−.47 (.17)
Kitchens-Stephens, 2005	Cognition	Short	8.27 (1.54)
Klak & Martin, 2003	Attitude	Short	.14 (.13)
Kohl, 2005	Cognition	Short	.82 (.32)
Kracht, 1998	Cognition	Short	.22 (.09)
Lal, 2010	Attitude	Short	1.24 (.23)
Landis et al., 1976	Cognitive	Short	.05 (.07)
Landis et al., 1985	Attitude	Short	.45 (.46)
Law, 1998	Attitude	Long	.19 (.24)
	Behavior	Long	.40 (.25)
	Cognition	Long	.49 (.25)
	Cognition	Short	1.19 (.26)
Lee et al., 2006	Cognition	Short	1.06 (.43)
Liberman et al., 2011			
Study 1	Reaction	Short	.51 (.2)
Study 2	Reaction	Short	.27 (.21)
Lichtenstein et al., 2008	Behavior	Long	.32 (.30)
	Cognition	Long	−.10 (.30)
	overall	Overall	.85 (.31)
Livosky et al., 2011	Attitude	Short	.53 (.20)
LoboPrabhu et al., 2000	Cognition	Short	.74 (.23)
Lopez-Humphreys, 2011	Cognition	Short	.14 (.20)
Luger, 2011	Attitude	Short	.14 (.20)
	Behavior	Short	.14 (.20)
	Cognition	Short	.14 (.20)
Madera et al., 2011	Attitude	Short	.20 (.10)
Majumdar et al., 2004	Attitude	Short	.41 (.12)
	Behavior	Short	.41 (.12)
	Cognition	Short	.41 (.12)
Majumdar et al., 1999	Behavior	Short	.52 (.33)
Mak & Buckingham, 2007	Behavior	Short	.48 (.29)
Manese et al., 2001	Attitude	Short	−.12 (.20)
	Cognition	Short	.79 (.23)
Manis, 2008	Attitude	Short	.08 (.20)
	Cognition	Short	.09 (.20)
	Behavior	Short	.11 (.20)
Mapp et al., 2007	Behavior	Short	.14 (.20)
Martin, 2006	Attitude	Short	.09 (.16)
	Cognition	Short	.55 (.16)
Martinez, 1995	Cognition	Short	.23 (.38)
Mausehund et al., 1995	Attitude	Short	.03 (.24)
May, 2010	Attitude	Short	.36 (.21)
McCleskey, 1991	Cognition	Short	.62 (.24)
	Attitude	Short	.46 (.24)
McCool et al., 2006	Attitude	Short	−.37 (.05)
Medina-Walpole et al., 2012	Attitude	Short	.20 (.10)
Middleton, 2002	Attitude	Short	.42 (.12)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Mio, 1989	Reaction	Short	.79 (.40)
Moffat & Tung, 2004	Attitude	Long	.49 (.15)
	Behavior	Long	.38 (.16)
Morin, 1974	Attitude	Short	.30 (.23)
Moss-Racusin et al., 2013	Behavior	Short	.86 (.20)
	Cognition	Short	.35 (.18)
	Attitude	Short	.21 (.18)
Moyer & Nath, 1998	Cognition	Short	.73 (.24)
Murphy et al., 2006	Attitude	Short	1.67 (.43)
	Behavior	Short	1.00 (.34)
	Cognition	Short	1.19 (.36)
Mysore, 2004	Attitude	Short	.52 (.25)
Neville & Furlong, 1994	Attitude	Long	.16 (.23)
	Behavior	Long	.55 (.23)
Neville et al., 1996	Attitude	Long	.48 (.21)
	Attitude	Short	.29 (.19)
	Behavior	Short	.69 (.18)
	Cognition	Long	.90 (.23)
	Cognition	Short	.69 (.18)
Nguyen et al., 2012	Behavior	Short	.27 (.11)
Nokes et al., 2005	Attitude	Short	.89 (.30)
	Cognition	Short	−1.22 (.34)
Paluck, 2009	Attitude	Long	.23 (.31)
Parker et al., 1998	Attitude	Short	.39 (.28)
Pedersen, 2010	Attitude	Long	.61 (.28)
Pedersen & Barlow, 2008	Attitude	Short	−.49 (.14)
Perry et al., 1998	Behavior	Short	−.12 (.35)
	Cognition	Short	.26 (.34)
Pettijohn & Walzer, 2008	Attitude	Short	.53 (.20)
Pilkington, 1993	Attitude	Short	.21 (.26)
	Behavior	Short	.46 (.27)
Preusser et al., 2011	Attitude	Short	.38 (.24)
	Behavior	Short	.41 (.24)
	Cognition	Short	.54 (.24)
	Reaction	Short	.25 (.24)
Probst, 2003	Attitude	Short	.30 (.21)
Randolph et al., 1977	Behavior	Short	.74 (.41)
Rehg et al., 2012	Attitude	Short	.27 (.16)
	Behavior	Short	.38 (.17)
	Cognition	Short	1.00 (.20)
Reinhardt, 1994	Attitude	Short	.26 (.11)
Remer, 2008	Attitude	Short	.10 (.13)
	Cognition	Short	.26 (.13)
Remmert, 1993	Cognition	Short	.16 (.20)
	Attitude	Short	.25 (.20)
Reynolds, 2010	Attitude	Short	.89 (.27)
	Cognition	Short	1.91 (.70)
Rhyne, 1973	Attitude	Short	.42 (.29)
	Attitude	Long	−.22 (.29)
Riggs et al., 2011	Attitude	Short	1.00 (.21)
	Behavior	Short	1.15 (.22)
	Cognition	Short	1.55 (.26)
Roberson et al., 2002			
Condition 1	Behavior	Short	.00 (.15)
Condition 2	Behavior	Short	.56 (.16)
Robinson & Bradley, 1997	Attitude	Short	.58 (.28)
	Cognition	Short	1.23 (.29)
Rogers-Sirin & Sirin, 2009	Cognition	Short	.29 (.10)
Rouh, 2001	Cognition	Short	1.20 (.19)
Rousey, 2010	Attitude	Short	.00 (.18)
	Behavior	Short	.09 (.18)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Rowell & Benshoff, 2008	Attitude	Short	.11 (.11)
Rudman et al., 2001			
Study 1	Attitude	Short	.95 (.36)
Study 2	Attitude	Short	.56 (.25)
Rudolph, 1989	Behavior	Short	.52 (.28)
	Attitude	Short	.48 (.28)
	Attitude	Long	.49 (.28)
Sakurai et al., 2010	Behavior	Long	.47 (.20)
Sanchez & Medkik, 2004	Behavior	Long	−.32 (.18)
	Cognition	Long	.2 (.18)
	Reaction	Long	.16 (.18)
Sanchez-Burks et al., 2007	Attitude	Long	.67 (.24)
	Behavior	Long	.74 (.24)
Sanner et al., 2010	Attitude	Short	.50 (.15)
Scher, 2008	Attitude	Short	.36 (.17)
	Cognition	Short	1.24 (.27)
Seguin, 2002	Attitude	Short	.02 (.06)
	Cognition	Short	.14 (.06)
Shergill, 1997	Behavior	Short	.21 (.32)
Shields et al., 2011	Reaction	Short	.64 (.19)
Soble et al., 2011	Attitude	Short	.13 (.18)
St. Clair & McKenry, 1999	Attitude	Short	.22 (.11)
	Attitude	Long	.39 (.15)
	Attitude	Short	.30 (.15)
Stebbins, 2005	Attitude	Short	.30 (.19)
Steed, 2010	Attitude	Short	−.07 (.27)
Steinfeldt & Wong, 2010	Attitude	Short	.38 (.30)
	Cognition	Short	.76 (.31)
Stella et al., 2007	Attitude	Short	.24 (.08)
Stewart et al., 2003	Attitude	Short	.53 (.24)
Stewart et al., 2010	Attitude	Long	.64 (.41)
Stowe, 2002	Cognition	Short	.16 (.38)
Tang et al., 2002	Cognition	Short	.12 (.08)
Tang et al., 2004	Attitude	Short	.36 (.28)
Taras et al., 2013	Attitude	Short	.04 (.02)
Taton, 2008	Attitude	Short	.28 (.25)
	Cognition	Short	.11 (.25)
Taylor-Ritzler et al., 2008	Attitude	Short	.11 (.06)
	Cognition	Short	.22 (.06)
Theis-Cole, 1995	Behavior	Short	.08 (.16)
	Cognition	Short	.16 (.16)
	Attitude	Short	.16 (.16)
Thomas & Cohn, 2006	Attitude	Long	.77 (.20)
	Attitude	Short	.61 (16)
Thomas, 2008	Behavior	Short	.29 (.22)
Timm, 1997	Attitude	Long	.51 (.24)
	Cognition	Long	.09 (.23)
Todd et al., 2011	Attitude	Short	.59 (.28)
Tolleson Knee, 1999	Attitude	Short	.11 (.17)
Tran et al., 1994	Attitude	Short	.13 (.13)
Turner, 1986	Cognition	Short	.20 (.35)
Van Soest, 1996	Attitude	Short	−.53 (.11)
	Behavior	Short	.24 (.11)
	Attitude	Short	−.34 (.18)
	Behavior	Short	−.29 (.18)
VonDras & Lor-Vang, 2004			
Sample a	Attitude	Short	.24 (.18)
Sample b	Attitude	Short	.28 (.16)
Sample c	Attitude	Short	.21 (.13)
Wade & Bernstein, 1991	Behavior	Short	.24 (.23)
Waight & Madera, 2011	Reaction	Short	−.02 (.23)

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B (*continued*)

Author, year	Outcome variable	Training effects	*g* (*SE*)
Wang, 1998	Attitude	Long	.99 (.27)
	Attitude	Short	1.09 (.27)
	Behavior	Long	1.11 (.27)
	Behavior	Short	.54 (.26)
	Cognition	Long	.55 (.26)
	Cognition	Short	.84 (.26)
Warring et al., 1998	Attitude	Short	.63 (.12)
	Behavior	Short	.55 (.12)
	Cognition	Short	1.28 (.15)
Weathersby, 2004	Attitude	Short	.57 (.12)
Wideman, 2005	Attitude	Short	.19 (.31)
Wiggins et al., 2007	Attitude	Short	.20 (.29)
	Cognition	Short	1.01 (.31)
Williams, 2005	Attitude	Short	.43 (.30)
	Behavior	Short	.18 (.30)
	Cognition	Short	.42 (.30)
Worchel & Mitchell, 1972	Attitude	Long	.38 (.20)
Wortman, 2002	Attitude	Short	.12 (.15)
Zarubin, 2008	Behavior	Long	.22 (.10)

Received June 25, 2014
Revision received June 7, 2016
Accepted June 13, 2016 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.