Journal of Diversity in Higher Education

Do Diversity Courses Improve College Student Outcomes? A Meta-Analysis

Nida Denson, Nicholas A. Bowman, Georgia Ovenden, KC Culver, and Joshua M. Holmes

Online First Publication, May 7, 2020. http://dx.doi.org/10.1037/dhe0000189

CITATION

Denson, N., Bowman, N. A., Ovenden, G., Culver, K., & Holmes, J. M. (2020, May 7). Do Diversity Courses Improve College Student Outcomes? A Meta-Analysis.*Journal of Diversity in Higher Education*. Advance online publication. http://dx.doi.org/10.1037/dhe0000189

American Psychological Association
NADOHE
National Association of Diversity Officers in Higher Education
Journal of Diversity in Higher Education

© 2020 National Association of Diversity Officers in Higher Education
ISSN: 1938-8926
http://dx.doi.org/10.1037/dhe0000189

Do Diversity Courses Improve College Student Outcomes? A Meta-Analysis

Nida Denson
Western Sydney University

Nicholas A. Bowman
University of Iowa

Georgia Ovenden
University of Melbourne

KC Culver
University of Southern California

Joshua M. Holmes
Suffolk University

Colleges and universities play a critical role in shaping intergroup dynamics in an era of increasing racial tensions in the United States. Diversity courses may serve as one important approach for preparing college students for participation in an equitable and just society, since this coursework holds a unique position at many institutions to expose college students to issues of difference and inequality. This study synthesizes research on the relationship between university/college instruction explicitly using the word course and the root divers` and student outcomes over the span of 25 years. Within a meta-analytic sample of 355 effect sizes, from 73 publications, and 47 distinct samples representing 116,092 undergraduate students the results indicate an overall small positive association between diversity coursework and various outcomes. Additional results highlighted the ways in which this relationship is moderated by various characteristics of the courses, outcome measures, and study design.

Keywords: diversity courses, curricular diversity, meta-analysis, college students

Supplemental materials: http://dx.doi.org/10.1037/dhe0000189.supp

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.

Over the past decade, there has been an increase in racial tensions and violence across the United States (e.g., Simon & Sidner, 2018). Not surprisingly, the racial tensions and racial politics of the nation are also reflected at colleges and universities. In 2016, there were 1,250 hate crimes (defined as crimes motivated by race, ethnicity, national origin, religion, sexual orientation, gender, or disability) reported on United States college campuses, an increase of 25% from just a year earlier (Bauman, 2018). While nearly all types of hate crimes (i.e., based on a person's disability, gender, race, religion, sexual orientation, national origin, and ethnicity) increased, offenses associated with racial bias were the most common and accounted for 40% of all hate crimes on college campuses (Bauman, 2018). These statistics and numerous highly publicized campus incidents over the last few years (e.g., Jaschik, 2017) illustrate how colleges continue to be racially charged environments today. For many students, their undergraduate institutions are significantly more racially diverse than their high schools and neighborhoods (Orfield, Ee, Frankenberg, & Siegel-Hawley, 2016), so college may be the first time they have been exposed to such a range of diverse ideas and people. One avenue through which higher education institutions can educate students about race and racial understanding is through diversity coursework. Nelson Laird, Engberg, and Hurtado (2005) define diversity courses as "courses that have content and methods of instruction that are inclusive of the diversity found in society" (p. 450).

As many scholars have noted, the overarching goal of diversity coursework in higher education is to equip students for participation in an equitable and just society (e.g., Banks, 2013; de Novais & Spencer, 2019). Increasingly, institutions across the country are realizing the importance of diversity courses in the curriculum. In a 2015 survey of 325 Chief Academic Officers at Association of American Colleges and Universities institutions (Hart Research Associates, 2016) regarding their general education curriculum, the vast majority (87%) offer diversity studies and experiences to their students, with one third (34%) of institutions implementing a requirement for all students to participate. Indeed, almost three quarters (73%) of institutions reported that "knowledge of diversity in the United States" is an important learning outcome for their students. However, exactly how institutions incorporate diversity

Nida Denson, School of Psychology, Western Sydney University; Nicholas A. Bowman, Department of Educational Policy and Leadership Studies, University of Iowa; Georgia Ovenden, Department of Social Work, University of Melbourne; KC Culver, Pullias Center for Higher Education, University of Southern California; Joshua M. Holmes, Institutional Research and Assessment, Suffolk University.

Correspondence concerning this article should be addressed to Nida Denson, School of Psychology, Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia. E-mail: n.denson@westernsydney.edu.au

courses varies considerably. For example, some institutions choose to implement a "diversity requirement," some focus on general education programs, and others choose to infuse their diversity courses or diversity content throughout the entire curriculum (which is less frequent and more difficult to implement; Nelson Laird, 2003; Sciame-Giesecke, Roden, & Parkison, 2009). What is also likely is that some institutions may not label diversity courses in this manner and/or may not provide an intentional integration of diversity content even though they may do so. According to the National Survey of Student Engagement (2019), 52% of senior students reported having taken courses that encouraged learning about other cultures either "quite a bit' or "very much." In addition, 50% and 65% of graduating students reported discussing issues of equity or privilege and respecting the expression of diverse ideas, respectively. Thus, it is clear that institutions consider diversity courses important, although to varying degrees. What is less clear is to what extent diversity courses affect student learning, growth, and development.

Reviews of Diversity Coursework and Student Outcomes

To date, two qualitative systematic reviews (Denson & Bowman, 2017; Engberg, 2004) and three quantitative meta-analyses (Bowman, 2010a, 2011; Denson, 2009) have examined the relationship between college diversity experiences and various student outcomes. Both Engberg (2004) and Denson (2009) focused on educational interventions that were designed to reduce racial bias; diversity coursework constitutes one such approach. Engberg's critical examination of educational interventions designed to reduce racial bias examined quantitative, qualitative, and mixed-methods studies, while Denson's meta-analysis synthesized quantitative studies on this topic. Bowman conducted two meta-analyses on college diversity experiences and general domains of student outcomes: cognitive growth (Bowman, 2010a) and civic engagement (Bowman, 2011). Overall, these reviews found that diversity courses were generally—but not always—associated with desired outcomes. These articles explored some moderators of this relationship, but diversity coursework constituted just one of several types of experiences, so the implications for course design and future research on coursework were not always clear.

The most recent and pertinent review was conducted by Denson and Bowman (2017); this article conducted a qualitative synthesis of research on college diversity courses and various outcomes. Their review consisted of 100 findings resulting from 92 primary studies that utilized qualitative, quantitative, and mixed methods approaches. Almost two thirds of these studies obtained a mix of mostly positive and nonsignificant outcomes of diversity courses, highlighting the substantial variability in the extant literature and the need for a more systematic understanding of these equivocal results. The authors concluded that the mixed findings are likely attributable to the outcome(s) examined, diversity course(s) examined, and study design, but the qualitative nature of this review (along with the inclusion of studies with qualitative and mixed-methods designs within the review) made it very difficult to draw specific conclusions. The most salient difference across studies pertained to the race/ethnicity of the sample: 62% of analyses with all-White samples yielded only positive and significant results, whereas only 29% of analyses with samples of all students of color identified only positive and significant results.

While informative, this review leaves some important unanswered questions. Specifically, does the relationship between diversity courses and student outcomes vary depending on the type of outcome examined? Are some types of diversity courses—or measures of diversity coursework—more strongly associated with student outcomes? To what degree is the size of this relationship related to sample characteristics or other aspects of study design? Thus, in the present study, we conducted a quantitative meta-analysis to synthesize the relationship between diversity courses and student outcomes and address the following three research questions: (a) does an overall relationship exist between college diversity courses and student outcomes?, (b) is there significant variation in this relationship across studies?, and (c) to what extent are study characteristics (i.e., type of outcome, type of diversity course, and study design) associated with the magnitude of this relationship?

The exclusive focus on diversity courses allowed this synthesis to examine specific aspects of coursework that have not appeared in prior reviews, such as the pedagogical approach within a specific course and the measurement of diversity coursework (e.g., number of courses vs. infusion of diversity throughout the curriculum).

We made the following hypotheses based on the results of the Denson and Bowman (2017) qualitative review as well as earlier systematic reviews that examined a narrower set of outcomes but a broad range of diversity experiences (Bowman, 2010a, 2011; Denson, 2009; Engberg, 2004). First, we expected to find a positive (albeit somewhat modest) relationship between diversity courses and student outcomes overall. Second, we expected to find considerable variation in this relationship across studies. Third, we anticipated that the variability across studies would be largely attributable to study characteristics. In particular, the relationship between diversity courses and student outcomes would be stronger when the outcome was attitudinal and/or related to diversity (e.g., intergroup prejudice). We expected that diversity courses with an interpersonal interaction component, such as through intergroup dialogue, would exhibit a stronger effect than diversity courses based on diverse content alone. Lastly, certain aspects of study design would also be a source of variation between studies; we hypothesized that studies which controlled for other college experiences (as well as other diversity experiences) in the statistical analyses would have smaller effects than studies that did not control for college experiences.

Method

Data Sources and Sampling Procedure

We conducted a search across several databases (ERIC, PsycINFO, ProQuest) covering 25 years of research (from January 1990 to December 2014). The beginning point for the search was employed to focus on studies that considered a reasonably recent version of a diversity course. The Denson (2009) meta-analysis that examined the effects of curricular and cocurricular diversity activities on students' racial bias included all available studies from the earliest date possible in each database search (i.e., ERIC from 1966–2006, PsycINFO from 1840–2006, and Dissertation

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.

Abstracts International from 1861–2006). The resulting search only showed one study prior to the 1990s, was more akin to a racial awareness workshop conducted over two weekends (with an all-White sample), and was an obvious outlier as compared to all the other studies which were published in 1993 and onward.

We utilized the following search terms:

> KEYWORDS = ("diverse" or "ethnic studies" or "multicultural studies" or "African American studies" or "Latin" studies" or "Asian American studies" or "women" studies" or "gender studies") AND "course" AND KEYWORDS = ("college student" or "undergraduate student" or "university student").

The search was limited to sources written in English only. We also searched for additional sources that were cited in the studies that we obtained. Lastly, we conducted a manual search of every article in the *Journal of Diversity in Higher Education*, *Journal of Higher Education*, *Research in Higher Education*, *Review of Higher Education*, and *Journal of College Student Development*. A study was used if it met all of the following inclusion criteria: (a) it was empirically grounded, (b) it investigated the relationship between at least one diversity course and at least one student outcome, (c) it reported quantitative data in sufficient detail for calculating effect sizes, (d) its participants were undergraduate students or were reporting about their previous undergraduate experience in the United States, and (e) it had a "publication date" from 1990 to 2014 (this criterion also includes unpublished work, such as the year in which a dissertation was completed or a conference presentation was given). We also considered whether to include service-learning courses, as they are arguably a form of diversity coursework. However, because service-learning courses are defined in terms of their use of a specific pedagogy (which cannot be differentiated from their content), these courses were excluded from the meta-analytic sample (the outcomes of service-learning have also been explored through prior meta-analyses; see Celio, Durlak, & Dymnicki, 2011; Conway, Amel, & Gerwien, 2009; Warren, 2012). In addition, the majority of studies often did not specify whether the diversity course was required or not, so we did not examine the differential effects for required versus nonrequired diversity courses. The sample for our meta-analysis consisted of 355 effect sizes from 73 publications, and 47 distinct "studies" (i.e., distinct samples) representing a total of 116,092 undergraduate students. An overview each study included in the meta-analysis is provided in the online supplemental materials.

Computing Effect Sizes

We utilized correlation coefficients as effect sizes. Since standardized regression coefficients (β) correspond reasonably well to correlation coefficients (*r*), we directly substituted beta coefficients for correlation coefficients if the beta coefficients were available (Peterson & Brown, 2005). The majority of the studies reported standardized betas. However, some studies reported partial correlations, means and standard deviations, unstandardized regression coefficients, *t* test analyses, analyses of variance, or analyses of covariance results. In the case of partial correlations, we substituted directly for *r*; for the bivariate analyses, we translated the findings into correlations. For studies that reported unstandardized coefficients, we computed standardized regression coefficients when reported data allowed the transformation.

For studies that reported multiple analyses that examined the same relationship (e.g., blocked hierarchical regression analyses), we used the coefficient(s) from the most fully identified model when an apparent mediator was not in the model. For studies that included interaction terms to examine moderating effects, we used the coefficient(s) from the block before the interaction terms were added. For studies that utilized structural equation modeling (SEM), we used the coefficient(s) from the most complete model without a mediator(s) when possible. For the SEM studies that contained models with mediators only, we used the coefficient(s) from the total effects (i.e., direct plus indirect effects). If only means and standard deviations were all that were available, we used them to calculate *d*, which was then converted into *r*. Because the product-moment correlation has some undesirable statistical properties, we utilized Fisher's Z_r transformation based on Hedges and Olkin (1985) recommendation. As recommended by Borenstein, Hedges, Higgins, and Rothstein (2009), we utilized Fisher's Z_r to perform the analyses, then converted the summary values back to correlations for presentation purposes. Thus, the coefficients in Tables 2–5 can be interpreted as standardized beta coefficients, similar to the Bowman (2011) meta-analysis.

Independent Variables and Coding

The study characteristics were our predictors or moderators of effect sizes. These consisted of the following: outcome type, diversity course type/measure, and study design characteristics (Table 1). Type of outcome was categorized in two ways: diversity-related outcome (0 = no; 1 = yes) and outcome type (1 = affective; 2 = cognitive; 3 = behavioral and behavioral intention). Examples of diversity-related outcomes are cultural awareness and promoting racial understanding, and examples of non-diversity-related outcomes are moral reasoning and need for cognition. If the study's outcome was a composite variable that included at least one variable that was diversity-related, the composite variable was classified as diversity-related. Outcome type was represented by two dummy variables, with behavioral and behavioral intention outcomes (e.g., interactions with diverse peers, orientation toward social/political activism) as the referent group. Examples of affective outcomes include modern sexism and liberal political views, and examples of cognitive outcomes are critical thinking skills and performance self-esteem.

A set of dummy variables was used to indicate the type or measure of diversity course: a single diversity course, number of courses (as a count or ordinal scale), curricular diversity composite (i.e., the extent to which diversity content was incorporated throughout students' coursework). Preliminary analyses showed no differences across disciplines/fields for single courses, and thus all the single courses (regardless of discipline) were combined into one single category (and was the referent group). Another dummy variable was used to indicate whether the course included a required discussion component (e.g., through intergroup dialogues).

Study design characteristics consisted of several variables, which varied across Level 1 and Level 2. For the study design characteristics at Level 1 (i.e., within distinct studies/samples), a dummy variable was used to indicate whether the study was published or not. Racial/ethnic composition of the sample was also included: Caucasian/White students only (referent group), students

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 1
Summary of Independent Variables, Coding, and Descriptive Statistics

Category of study characteristic	Predictor variables	*N*	Min	Max	*M*	*SD*
Outcome type (Level 1)	Behavioral and behavioral intent (referent group)	355	0	1	.20	.40
	Affective	355	0	1	.65	.48
	Cognitive	355	0	1	.15	.36
	Diversity-related outcome (vs. non-diversity-related outcome)	355	0	1	.60	.49
Diversity course type (Level 1)	Single diversity course (referent group)	355	0	1	.73	.45
	Number of courses	355	0	1	.19	.39
	Curricular diversity composite	355	0	1	.09	.28
	Course includes a required discussion component (vs. course did not include a required discussion component)	355	0	1	.36	.48
Study design characteristics (Level 1)	Published (vs. unpublished)	355	0	1	.85	.35
	All Caucasian/White sample (reference group)	355	0	1	.12	.32
	Mixed racial/ethnic sample	355	0	1	.81	.39
	All students of color sample	355	0	1	.07	.26
Study design characteristics (Level 2)	Cross-sectional outcomes (referent group)	47	0	1	.40	.50
	Self-reported gains	47	0	1	.04	.20
	Longitudinal gains	47	0	1	.55	.50
	Multiple institutions within sample (vs. single institution)	47	0	1	.38	.49
	Included multiple diversity experiences in the model (vs. did not include multiple diversity experiences)	47	0	1	.66	.48
	Included other college experiences in the model (vs. did not include other college experiences)	47	0	1	.77	.43
	Various years (referent group)	47	0	1	.43	.50
	Freshman year	47	0	1	.15	.36
	Sophomore year	47	0	1	.09	.28
	Junior year	47	0	1	.06	.25
	Senior year	47	0	1	.21	.41
	After college	47	0	1	.06	.25

of color only, or a mixed-race sample. For the study design characteristics at Level 2 (i.e., across distinct studies/samples), one study design characteristic reflected the way in which the dependent variable was assessed: cross-sectional design that used students' ratings of their current attributes with no pretest (which served as the referent group), self-reported gains (i.e., gains measured reflectively by students), or longitudinal measurement of gains. Other study design characteristics were represented by dummy variables: whether the sample was multi-institutional (vs. single institution); whether the study examined multiple diversity experiences within the same statistical model (0 = no; 1 = yes); and whether the study controlled for other college experiences (0 = no; 1 = yes). We also included the year in college in which the final data collection occurred: 1 = first year; 2 = sophomore year; 3 = junior year; 4 = senior year; 5 = postcollege; and 6 = various years (referent group); according to our preliminary analyses, other approaches for coding this construct yielded the same substantive conclusions.

Three of the authors worked together to create the coding categories and coding sheet. As a preliminary stage, two authors individually reviewed several publications on two separate occasions, and then they discussed their results with each other and a third author. After they were confident about the coding procedure, the two authors each separately coded 10% of the publications. The interrater reliability was very high (with a mean Cohen's kappa across categories of .95), so the remaining studies were coded individually. The first author randomly selected five publications to ensure that the coding was done correctly, which they were. The first author also coded the effect sizes for all the publications.

Analyses

We utilized hierarchical linear modeling (HLM; Version 7.01) to conduct the meta-analysis due to the multilevel nature of the studies, as participants are nested within studies, and the relevant effects can occur both within and across studies (Raudenbush & Bryk, 2002; Raudenbush, Bryk, & Congdon, 2004). In addition, one of the main strengths of multilevel modeling is that it addresses the issue of nonindependence of observations by modeling each study as a Level 2 group. Consistent with other meta-analyses in higher education, "a study consists of a set of data collected under a single research plan from a designated sample of participants" (Lipsey & Wilson, 2001, p. 76). HLM analyses weight each effect size by the inverse of the study's squared standard error. Thus, larger weights are assigned to studies with larger sample sizes and smaller errors. Consistent with prevailing recommendations (e.g., Lipsey & Wilson, 2001), we use the term "study" to refer to data collected under a single research project using the same sample of participants, thereby distinguishing among samples instead of publications. Thus, while a study inmost cases does indeed correspond to a single publication, there are instances in which multiple publications arise from one study. For example, there were several articles that utilized data from the Preparing College Students for a Diverse Democracy study (i.e., Engberg, 2007; Engberg & Hurtado, 2011; Engberg, Hurtado, & Smith, 2007; Hurtado, 2005). In total, 73 publications corresponding to 47 distinct studies (i.e., samples) were modeled at Level 2.

HLM is ideal for meta-analyses in which effect sizes are nested within studies (i.e., each study reports several effect sizes), and moderators can vary both within studies and between studies. For

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.

example, a single study (e.g., the Diverse Democracy study) may have examined multiple outcome variables, multiple curricular diversity activities, and may have resulted in multiple publications (e.g., dissertation, journal article) that utilized different samples (e.g., mixed sample, students of color only). These attributes would be considered within-study effects (i.e., at Level 1). There are also other phenomena that generally vary across studies (i.e., distinct samples), such as how the outcome variable was assessed (e.g., cross-sectional, self-reported gains, longitudinal) and whether the sample consisted of students from a single institution or multiple institutions. These attributes would be considered between-study effects (i.e., at Level 2).

First, we explored descriptive statistics for the unweighted effect sizes. Next, we examined an unconditional HLM model with the effect size as the dependent variable and no independent variables. This unconditional model produces a weighted estimate of the overall effect size across all the studies, as well as a homogeneity analysis that tests whether there is more variation across studies than can be expected by chance (i.e., due to measurement error or random differences across studies). If the homogeneity test is significant, then a substantial amount of heterogeneity in the effect sizes may be explained by study characteristics. Thus, we included predictor (or moderator) variables in subsequent models that may explain some of the heterogeneity in effect sizes. A final full model that included significant predictors from the individual analyses was also created. All the dichotomous independent variables were uncentered, so the intercept represents the effect size for the referent group(s; Raudenbush & Bryk, 2002). By keeping the dichotomous predictors uncentered, we can determine simultaneously whether: (a) the effect size for the referent group differs significantly from zero, and (b) whether the predictor variables (i.e., study characteristics) are significantly related to the effect size.

Results

Descriptive Statistics and Preliminary Analyses

When exploring descriptive statistics for the 355 unweighted effect sizes (i.e., with each individual effect size given equal weight), there is an overall positive relationship between diversity courses and student outcomes ($M = .081$). Based on effect size recommendations for college impact research from Mayhew et al. (2016), this average effect size for the standardized regression coefficients is between small ($\beta = .06$) and medium ($\beta = .12$). Out of the 355 unweighted effect sizes, 36 effect sizes (10%) are negative in direction, 202 (57%) are positive but less than .10, 82 (23%) are from .10 to less than .20, 24 (7%) are from .20 to less than .30, and 11 (3%) effect sizes are at least .30. We created a funnel plot between sample size and effect size to visually assess the possible presence of publication bias (Borenstein et al., 2009; Lipsey & Wilson, 2001). The plot was in the shape of a funnel, suggesting that publication bias was not a concern.

HLM Analyses

The HLM analyses accounted for the multilevel structure of the data and weighted the effect sizes as a function of the study's variance.[1] The unconditional HLM analysis indicates that diversity courses overall are significantly and positively related to student outcomes ($B = .094$, $SE = .013$, $p < .001$). The 95% confidence interval for this effect size estimate is above zero: [.069, .119]. Although this review sought to include all studies on the topic, it is possible that some researchers chose not to publish or present nonsignificant findings or that nonsignificant results were not accepted for publication or presentation; this phenomenon is sometimes referred to as the "file drawer problem." Applying trim and fill procedures (Duval & Tweedie, 2000) to a funnel plot of the standard errors by effect sizes for the 47 studies (i.e., independent samples), 13 unpublished studies/samples were projected to be missing from the left side of the funnel plot. This suggests that there is some publication bias present, and that the mean effect size of .094 is an overestimate of the true effects of diversity courses. However, the mean effect size is still positive, although smaller in magnitude ($B = .049$, $SE = .016$, $p = .002$) and the 95% confidence interval is still above zero: [.018, .080]. In addition, a Fail-safe N was computed to determine how many studies with an average effect size of zero would need to be added to the current sample to make the overall effect size nonsignificant (see Rosenthal, 1979). The Fail-safe N for this meta-analysis was 905 studies, which suggests that the present effect size estimate (based on 47 independent studies or samples) reflects a legitimate positive (albeit small) relationship between diversity coursework and college student outcomes. More importantly, substantial heterogeneity in effect sizes exists across studies, $\chi^2(46) = 2421.118$, $p < .001$, suggesting that further models are needed to explain this variance. In addition, the I^2 (measure of the proportion of the observed variance that reflects real differences in effect size) for the overall effect was 91.66, and is considered to be high (Higgins, Thompson, Deeks, & Altman, 2003). This means that 91.66% of the dispersion in diversity course effect sizes reflect real differences in effect size, and that 8.34% was due to random error.

To avoid problems with the relatively small sample size, we conducted several HLM analyses with a limited number of predictors to determine which variables should be entered into the final model. The first moderation analysis examined whether the magnitude of the effect size depended on the type of outcome examined. Table 2 shows that diversity courses are positively related to behavioral or behavioral intention outcomes that are not diversity related, as indicated by the results for the intercept ($B = .062$, $p < .001$), a small effect according to Mayhew et al. (2016). That is, the intercept represents the average effect size for studies that are in the referent group for all predictors (i.e., the values of the independent variables all equal 0). This effect size is significantly larger for diversity-related outcomes than non-diversity-related outcomes ($B = .034$, $p < .05$), but no significant differences exist between outcomes that are behavioral versus those that are affective ($B = .011$, $p = .159$) or cognitive ($B = .012$, $p = .159$). To examine whether diversity courses are positively related to student outcomes regardless of outcome type (i.e., affective, cognitive, or behavioral/behavioral intentions), additional analyses were conducted using each outcome type as the intercept and then

[1] HLM weights each effect size by the inverse of the study's squared standard error, which assigns larger weights to studies with larger samples (smaller errors) and smaller weights to studies with smaller sample sizes (larger errors).

Table 2
Unstandardized Coefficients for Hierarchical Linear Modeling Analyses of Outcome Type and Course Type/Measure Predicting Effect Size

Independent variable	Coefficient	*SE*	*df*	*t* ratio
Intercept	.062```	.014	46	4.566
Affective outcome (vs. behavioral or behavioral intentions)	.011	.008	305	1.411
Cognitive outcome (vs. behavioral or behavioral intentions)	.012-	.008	305	1.451
Diversity-related outcome (vs. non-diversity-related outcome)	.034	.015	305	2.292
Intercept	.082```	.015	46	5.317
Number of courses (vs. single diversity course)	.006--	.013	305	.478
Curricular diversity composite (vs. single diversity course)	.052	.019	305	2.813
Course included a structured discussion component (vs. course that did not include a structured discussion component)	.014	.013	305	1.073

Note. Horizontal lines distinguish between separate analyses; in other words, predictors for outcome type and course type were examined in separate models.
` $p < .05$. `` $p < .01$. ``` $p < .001$.

with the diversity-related outcome variable excluded from the models. In all analyses, the intercept was significantly greater than zero (*B*s ≥ .082, *p*s < .001), suggesting that diversity courses are positively associated with each of the three types of outcomes, although a small effect again (according to Mayhew et al., 2016).

The second analysis explored whether effect size depended on diversity course type or measure. Table 3 shows that taking a single diversity course (regardless of its discipline) is positively related to student outcomes, as indicated by the intercept ($B = .082$, $p < .001$). There are no significant differences, however, between measures of a single diversity course and the number of diversity courses taken ($B = .006$, $p = .633$) or between single diversity courses and courses that included a structured discussion component ($B = .014$, $p = .284$). In contrast, studies that used a curricular diversity composite measure have significantly larger effect sizes than do those that include a single diversity course ($B = .052$, $p < .01$). Similar to outcome type, additional analyses were also conducted on diversity course type/measure. The intercept was significantly greater than zero across all analyses (*B*s ≥ .071, *p*s < .001), suggesting that all types and measures of diversity coursework are positively associated with student outcomes.

A number of analyses were conducted to examine study design characteristics. Some study design variables were modeled at Level 1, since considerable within-study variation exists for different-race samples (i.e., numerous studies conducted subgroup analyses), and some studies had articles that are either published or unpublished. In contrast, there was relatively less variation for the Level 2 characteristics (e.g., single or multiple institutions within the study sample). In terms of Level 1 characteristics, Table 3 shows that unpublished studies with a sample that consists only of White students are positively related to student outcomes ($B = .076$, $p < .001$). However, the magnitude of the effect sizes does not depend on the racial/ethnic composition of the samples. In other words, there are no significant differences in effect sizes between studies that examined only students of color versus only White students ($B = -.013$, $p = .220$) or between studies that utilized an all-White student sample versus a mixed racial/ethnic sample ($B = -.013$, $p = .253$). (Supplemental analyses that explored the robustness of this result, such as only considering studies that included samples of both students of color and White students as well as separating students of color into several distinct groups by race, also yielded nonsignificant findings.) In terms of Level 2 study design characteristics, Table 4 provides the results for several distinct analyses. Studies that measure student outcomes with across-sectional indicator of outcomes have a significantly larger average effect size than longitudinal studies ($B = -.082$, $p < .001$) and studies in which students provided self-reported gains ($B = -.092$, $p < .001$). Studies that controlled for other college experiences ($B = -.088$, $p < .001$) and other diversity experiences ($B = -.061$, $p < .001$) also have significantly smaller effect sizes than studies that did not control for either type of experience. However, no significant differences are evident for single- versus multiinstitutional studies or for students' year in college.

Table 3
Unstandardized Coefficients for Hierarchical Linear Modeling Analyses of Other Level1 Study Characteristics Predicting Effect Size

Independent variable	Coefficient	*SE*	*df*	*t* ratio
Intercept	.076```	.016	46	4.810
Published (vs. unpublished)	.022`	.011	307	2.102
Intercept	.105```	.015	46	7.019
Mixed racial/ethnic sample (vs. all Caucasian/White sample)	−.013	.012	306	−1.146
Only students of color in the sample (vs. all Caucasian/White sample)	−.013	.010	306	−1.230

Note. Horizontal lines distinguish between separate analyses.
` $p < .05$. ``` $p < .001$.

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 4
Unstandardized Coefficients for Hierarchical Linear Modeling Analyses of Level 2 Study Characteristics Predicting Effect Size

Independent variable	Coefficient	*SE*	*df*	*t* ratio
Intercept	.144```	.016	44	8.962
Self-reported gains (vs. cross-sectional)	−.092```	.018	44	−5.205
Longitudinal (vs. cross-sectional)	−.082	.023	44	−3.573
Intercept	.109```	.018	45	6.146
Multiple institutions within sample (vs. single institution)	−.036	.024	45	−1.491
Intercept	.135```	.025	45	5.348
Included multiple diversity experiences in the model (vs. did not include multiple diversity experiences)	−.061`	.028	45	−2.150
Intercept	.162```	.015	45	10.678
Included other college experiences in the model (vs. did not include other college experiences)	−.088	.021	45	−4.282
Intercept	.114```	.023	41	5.086
First year (vs. various years in college)	−.018	.048	41	−.384
Second year (vs. various years in college)	−.016	.037	41	−.427
Third year (vs. various years in college)	−.019	.029	41	−.667
Fourth year (vs. various years in college)	−.052	.030	41	−1.736
Postcollege (vs. various years in college)	−.037	.044	41	−.844

Note. Horizontal lines distinguish among separate analyses.
` $p < .05$. ``` $p < .001$.

Table 5 presents the results of the full HLM analysis, which included all the significant predictors in the previous models (i.e., diversity-related outcomes, curricular diversity composite, and publication status at Level 1; self-reported gains, longitudinal outcomes, and controlling for other diversity experiences and other college experiences at Level 2). The number of courses and courses that included a structured discussion component were also included in the final model so that single courses—regardless of discipline—would be the referent group. At Level 1, the same predictors that were significant in previous models were also significant in the final combined model. That is, studies that examined a diversity-related outcome had larger effect sizes than studies that examined a nondiversity outcome ($B = .035$, $p < .05$), and studies that used a curricular diversity composite had larger effect sizes than studies that examined a single diversity course ($B = .058$, $p < .001$). At Level 2, studies that used a longitudinal design had smaller effect sizes than those that used cross-sectional measures of student outcomes ($B = -.069$, $p < .05$). However, self-reported gains ($B = -.049$, $p = .077$), other diversity experiences ($B = -.011$, $p = .823$), and other college experiences ($B = -.029$, $p = .471$) are no longer significant in the final model. All of the significant predictors in the model would be considered to be small effects (Mayhew et al., 2016).

Discussion

The present meta-analysis synthesized studies that examined the relationship between university/college instruction explicitly using the words course and the root divers` and various student outcomes over the span of 25 years. The results indicate that these courses are positively related to student outcomes, however, this positive effect is somewhat small in magnitude. While 90% of the 355 effect sizes (and 91% of the 47 effect sizes at the study level) in this meta-analysis are greater than zero, there is significant variation in effect sizes. Within the multilevel analyses (taking into account that participants are nested within studies), this positive association is significant across outcome type (affective, cognitive, and behavioral/behavioral intentions) as well as diversity course

Table 5
Unstandardized Coefficients for Hierarchical Linear Modeling Analysis of Level 1 and Level 2 Independent Variables Predicting Effect Size (Full Model)

Independent variable	Coefficient	*SE*	*df*	*t* ratio
Intercept	.119```	.016	42	7.397
Self-reported gains (vs. cross-sectional)	−.049`	.027	42	−1.811
Longitudinal (vs. cross-sectional)	−.069`	.026	42	−2.627
Included multiple diversity experiences in the model (vs. did not include multiple diversity experiences)	−.011	.047	42	−.225
Included other college experiences in the model (vs. did not include other college experiences)	−.029	.039	42	−.727
Diversity-related outcome (vs. non-diversity-related outcome)	.035	.014	303	2.546
Number of courses (vs. single diversity course)	−.003``	.013	303	−.211
Curricular diversity composite (vs. single diversity course)	.058	.017	303	3.369
Course included a structured discussion component (vs. course that did not include a structured discussion component)	.012	.012	303	1.058
Published (vs. unpublished)	.014	.012	303	1.245

` $p < .05$. ``` $p < .001$.

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.

type/measure (single course, number of courses, curricular diversity composite, and courses that included a required discussion component). The significant results also hold for the most rigorous studies that use longitudinal measurements of student outcomes, multiinstitutional samples, and appropriate control variables in the statistical models.

Of course, this overall result masks considerable variation in effect sizes within and across studies. As a result, moderator analyses were conducted to examine whether certain study characteristics would be able to explain some of this heterogeneity. The most robust finding regarding the nature of diversity courses is that overall curricular diversity (i.e., the extent to which diversity content was incorporated throughout students' coursework) is more positively related to student outcomes than is taking a single diversity course; this result even persists within the full model that accounts for other study characteristics. This pattern does not simply mean that adding more diversity coursework is always beneficial, since the number of diversity courses is not more strongly associated with student outcomes than simply taking a single diversity course. Instead, it appears that infusing diversity throughout the curriculum—which could occur via some "formal" diversity coursework and other courses that contain diversity content—may constitute the most effective strategy for bolstering student growth. This infusion approach has the potential benefit of reducing the resistance to learning about diversity in formally labeled diversity courses or requirements (e.g., Bowman, 2009) and by providing repeated exposure to information that may cause students to reconsider their beliefs and worldviews (e.g., Crisp & Turner, 2011).

Having a structured discussion component within a diversity course, such as through intergroup dialogue, is not more strongly associated with desired outcomes. Supplemental analyses that focused solely on intergroup dialogue (and excluded other discussion-based approaches) also do not yield significant results. This finding is surprising from both a theoretical and empirical perspective. Various theoretical perspectives highlight the role that interpersonal interactions play in improving intergroup attitudes and other outcomes (Allport, 1954; Crisp & Turner, 2011; Gurin, Dey, Hurtado, & Gurin, 2002; Pettigrew, 1998). Moreover, in her meta-analytic review of curricular and cocurricular diversity and racial bias, Denson (2009) found that interventions involving student contact were associated with greater bias reductions. The current meta-analysis examined a broad range of outcomes (not just prejudice or bias), which could at least partially explain why this study did not find a larger effect size for courses with a structured discussion component.

The potential impact of diversity courses appears to depend upon the type of outcome, since diversity-related outcomes (e.g., intergroup attitudes) have significantly larger relationships than other outcomes (e.g., general civic attitudes). This pattern makes sense, given that the content of diversity courses focuses directly on these outcomes, so one should expect these outcomes to exhibit larger effects. In contrast, the differences among affective, cognitive, and behavioral outcomes are nonsignificant. This finding is inconsistent with Ajzen's (1985, 1991) framework, which suggested that behaviors and behavioral intentions are relatively distal (or indirect) outcomes and should therefore be associated with smaller effect sizes than attitudinal or cognitive outcomes. In this meta-analytic sample, approximately two thirds of the studies examined attitudinal outcomes, so the differences between outcome types might have been significant at the traditional $p < .05$ criterion if more studies examining cognitive, behavioral, and behavioral intentions had been available.

Other findings pertained to methodological attributes of the research studies. In the initial analyses and full model, studies that used across-sectional assessment of students' current outcomes yield significantly larger relationships than those that measured change in outcomes longitudinally; the average effect size is about twice as large for cross-sectional as for longitudinal results. It appears that cross-sectional approaches may overestimate the impact of diversity coursework, especially since the pretest is often the most important predictor when accounting for self-selection into college experiences and other forms of engagement (Pascarella, Salisbury, & Blaich, 2013; Steiner, Cook, Shadish, & Clark, 2010). Interestingly, self-reported gains and longitudinal gains are associated with similar average effect sizes in the present study, despite the fact that self-reported gains are often a poor proxy for actual student growth (Herzog & Bowman, 2011; Porter, 2013). In addition, previous meta-analyses of college diversity experiences and student outcomes also identified no significant differences between longitudinal growth and cross-sectional outcome assessments (Bowman, 2010a, 2011). However, self-selection may be a larger problem for diversity coursework, since students choose to participate as a function of both general interest in electives and through their undergraduate major, which may range dramatically in diversity content.

In addition, other study characteristics are not significant moderators, especially when accounting for other variables. Year in college is unrelated to the magnitude of the effect size, which suggests that diversity coursework may be equally effective regardless of when it occurs during students' undergraduate experience. Moreover, in the initial analyses of this study, the presence of multiple diversity courses or other college experiences in the model predicts smaller associations between diversity coursework and student outcomes. This pattern is consistent with the use of control variables to provide a more conservative estimate of the relationship as well as findings from prior research syntheses (Bowman, 2010a, 2011; Denson, 2009). Published studies also have a modestly greater average effect size than unpublished studies, which is consistent with the apparent presence of publication bias throughout educational research (Polanin, Tanner-Smith, & Hennessy, 2016). However, all of these relationships decrease in magnitude and become nonsignificant in the full model that accounts for other between- and within-study predictors. It is possible that these initial findings were themselves confounded with other variables or that the somewhat limited number of studies at Level 2 reduced the variance that could be explained simultaneously with several between-study predictors.

Perhaps the most surprising finding is the lack of difference between samples that consist entirely of students of color, entirely of White students, or a mix of these groups. This result was consistently replicated when conducting follow-up sensitivity analyses, such as removing all mixed-race samples from the model. The greater efficacy of diversity courses for majority students—and particularly White students—has been widely assumed and even supported by Denson and Bowman's (2017) qualitative systematic review. They conducted a "vote count" procedure for identifying the pattern of significant results by

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.

racial/ethnic group. In their analysis, most studies of White samples (62%) found uniformly positive and significant results, which contrasted notably with the corresponding proportions for samples of African Americans (35%), Asian Americans (27%), Latinos/as (24%), and various students of color (30%). The problem with this qualitative review approach is that statistical significance is a product not only of effect size, but also of sample size. Students of color are less represented numerically on college campuses, so they are also less represented in research on diversity coursework. The present quantitative meta-analysis disentangles sample size and effect size, thereby showing that the effect size for diversity coursework is similar for students from minority and majority racial identities.

Finally, this article also provides important implications for examining conditional effects. As Seifert, Bowman, Wolniak, Rockenbach, and Mayhew (2017) point out, higher education researchers frequently conduct separate analyses for two or more subgroups (e.g., by race/ethnicity) without first examining whether the relationships differ significantly from one another. The fact that the result is significant for one group but not another is often erroneously interpreted as meaning that an educational practice is effective only for group(s) that have significant coefficients, even when the sample sizes differ substantially across groups. The present study illustrates the danger of that approach from a research synthesis perspective, since the average effect sizes for diversity coursework among students of color and White students are very similar, whereas Denson and Bowman (2017) found that results are about twice as likely to be statistically significant among White students as among students of color. Therefore, individual studies must directly assess whether relationships differ significantly across groups, which can occur either by entering an interaction term into a full-sample analysis or by conducting subgroup analyses and then testing for significant differences between coefficients (see Cohen, Cohen, West, & Aiken, 2013). As a good example of this approach, Gurin, Nagda, and Zúñiga (2013) conducted a number of moderation analyses (using interaction terms) to examine whether intergroup dialogues were more or less effective depending on the student's identity (male/White students vs. female/students of color). They found similar effects of intergroup dialogues on 20 of the 24 outcomes; of the four showing differential effects, these were larger for the groups with more societal privilege (i.e., male/White students) as compared to students with less privilege (i.e., female/students of color). Their findings are consistent with the present meta-analytic findings that suggest positive effects for all groups and generally no difference across groups.

In summary, this meta-analysis has demonstrated that diversity courses have a small, but positive effect on college student outcomes. However, this meta-analysis has also shown how various study characteristics may influence the results of those studies that are evaluating the effectiveness of such diversity courses, and points to suggestions for future research. For example, the Alimo (2012) study had a relatively large effect size as compared to the other studies. This was a multiinstitution study that utilized an experimental design with stratified random assignment. Potential participants were students who expressed interest in enrolling in a race/ethnicity dialogue, and then students were randomly assigned to one of two groups: an experimental group and await-list control group. The course was part of the Multi-University Intergroup Dialogue Research Project, which has an explicit focus on investigating the impacts of intergroup dialogues. While the study spanned across nine institutions, extra precautions were taken to ensure that the individual intergroup dialogue programs, and implementation of those programs, were as identical as possible across all nine campuses minimizing threats to validity (e.g., nearly identical curricula, activities, readings, syllabi, and common goals for training dialogue facilitators). Further care was taken in choosing the optimal size of the experimental and control groups (i.e., 16 students consisting of approximately four women of color, four White women, four men of color, and four White men). Thus, the Alimo (2012) study is an excellent example of research that has successfully overcome some of the obvious threats to selection bias in examining the effectiveness of diversity courses.

Limitations

There were some limitations that should be noted. First, the search terms were somewhat limited in scope and may have not identified all the relevant literature, since some synonyms were not included in the search string. For example, "course" was used which omitted synonyms such as class, seminar, or coursework. Second, there may be some missing studies that have not been included in the current meta-analysis, such as unpublished research that have not appeared as a conference article or dissertation. Since only three electronic databases were included in this study and other strategies for obtaining fugitive or gray literature were not conducted (e.g., contacting authors to solicit additional/unpublished articles, posting on listservs), it is possible that research may have been inadvertently excluded. Third, the majority of the studies only provided limited information about the pedagogy of these courses. For example, the extent to which privilege (among other issues) was discussed directly is unknown. On a related note, for courses that included a discussion component, it was unclear how much (or how little) discussion was actually incorporated into the course. Fourth, it is likely that some of the original studies that examined the effectiveness of a single diversity course may have conflated its effect with those of multiple courses (e.g., a student may have already taken another diversity course prior to the current one). Lastly, it was impossible to explore other relevant student characteristics beyond race, such as students' prior diversity experiences before college, that may have affected the relationship between diversity courses and student outcomes.

Conclusion

Higher education institutions around the country can and should play a critical role in shaping these intergroup dynamics. Past research has shown that interpersonal intergroup interactions and friendships, including those that occur during college, predict improved intergroup attitudes (see Davies, Tropp, Aron, Pettigrew, & Wright, 2011; Pettigrew & Tropp, 2011). Colleges and universities have some limited control over the quantity and quality of intergroup interactions, but they can certainly require students to take coursework that focuses on diversity as part of their general education. Thus, diversity coursework holds a unique position on many campuses as a shared experience with issues of difference and diversity. However, creating a common form of curricular engagement also comes with significant challenges, since such

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.

courses seek to promote learning and growth among students who range from being virtually unaware about issues of inequality to other students who have spent their whole lives confronted by it.

This synthesis provides not only support for the overall educational efficacy of diversity coursework, but also important insights for how diversity learning might be best implemented into the college curriculum. The strongest relationship with desired outcomes occurred for the inclusion of diversity throughout students' coursework, as opposed to taking one or more specific diversity-focused courses. Therefore, the presence of a general education requirement constitutes an important step, but it is certainly not sufficient for maximizing the benefits of curricular diversity. Students in some undergraduate majors, such as sociology or English, may encounter substantial diversity-related content throughout their undergraduate careers, but this emphasis contrasts notably with the experiences of science or engineering majors at many institutions. A challenge for faculty and administrators is how to promote the inclusion of diverse course material when some faculty will be reluctant or ill-prepared to do so. Campus teaching and learning centers could play a key role in preparing instructors appropriately. Moreover, given that year in college is unrelated to the apparent efficacy of this approach, students should be provided with opportunities to engage with curricular diversity both earlier and later in their course of study.

Given the debate on whether diversity courses have any impact on student growth (e.g., Herzog, 2010), this meta-analysis lends support for the consistent, positive link between college diversity coursework and various student outcomes and therefore the continued use of this approach in higher education. However, further research is now needed to pinpoint exactly what about diversity courses specifically (e.g., contextual and pedagogical practices) influence student learning and behavior. As Hurtado, Alvarez, Guillermo-Wann, Cuellar, and Arellano (2012) hypothesized in their model for diverse learning environments, the central features of effective curricular and cocurricular experiences should focus on the whole educational package, which includes "who we teach (student identities), who teaches (instructor identities), what is taught (content), and how it is taught (pedagogies/teaching methods)" (p. 49). The present analysis also suggests that longitudinal designs and the inclusion of appropriate control variables will likely provide more accurate estimates of the true relationships, thereby illustrating the need for rigorous designs in future research. Only with this nuanced understanding can we effectively develop diversity courses that will provide the strongest possible benefits for all students.

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.

References

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

Ajzen, I. (1985). *From intentions to actions: A theory ofplanned behavior* (pp. 11–39). Berlin/Heidelberg, Germany: Springer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179–211. http://dx.doi.org/10.1016/0749-5978(91)90020-T

*Alimo, C. J. (2012). From dialogue to action: The impact of cross-race intergroup dialogue on the development of white college students as racial allies. *Equity & Excellence in Education, 45,* 36–59. http://dx.doi.org/10.1080/10665684.2012.643182

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

*Antonio, A. L. (2001). Diversity and the influence of friendship groups in college. *Review of Higher Education, 25,* 63–89. http://dx.doi.org/10.1353/rhe.2001.0013

*Antony, J. (1993, November). *Can we all get along? How college impacts students' sense of the importance of promoting racial understanding*. Paper presented at the annual meeting of Association of the Study of Higher Education, Pittsburgh, PA.

Banks, J. A. (2013). *An introduction to multicultural education* (5th ed.). Boston, MA: Pearson.

Bauman, D. (2018, February .16). After 2016 election, campus hate crimes seemed to jump. Here's what the data tell us. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/After-2016-Election-Campus/242577

*Bolen, K. J. (2010). *Assessing the impact of diversity courses on student-faculty interactions, critical thinking and social engagement* (Unpublished doctoral dissertation). University of Southern California, Los Angeles, CA.

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

*Bowman, N. A. (2009). College diversity courses and cognitive development among students from privileged and marginalized groups. *Journal of Diversity in Higher Education, 2,* 182–194. http://dx.doi.org/10.1037/a0016639

Bowman, N. A. (2010a). College diversity experiences and cognitive development: A meta-analysis. *Review of Educational Research, 80,* 4–33. http://dx.doi.org/10.3102/0034654309352495

*Bowman, N. A. (2010b). The development of psychological well-being among first-year college students. *Journal of College Student Development, 51,* 180–200. http://dx.doi.org/10.1353/csd.0.0118

*Bowman, N. A. (2010c). Disequilibrium and resolution: The nonlinear effects of diversity courses on well-being and orientations toward diversity. *Review of Higher Education: Journal of the Association for the Study of Higher Education, 33,* 543–568. http://dx.doi.org/10.1353/rhe.0.0172

Bowman, N. A. (2011). Promoting participation in a diverse democracy a meta-analysis of college diversity experiences and civic engagement. *Review of Educational Research, 81,* 29–68. http://dx.doi.org/10.3102/0034654310383047

*Bowman, N. A. (2012). Promoting sustained engagement with diversity: The reciprocal relationships between informal and formal college diversity experiences. *Review of Higher Education: Journal of the Association for the Study of Higher Education, 36,* 1–24. http://dx.doi.org/10.1353/rhe.2012.0057

*Bowman, N. A., Brandenberger, J. W., Hill, P. L., & Lapsley, D. K. (2011). The long-term effects of college diversity experiences: Well-being and social concerns 13 years after graduation. *Journal of College Student Development, 52,* 729–739. http://dx.doi.org/10.1353/csd.2011.0075

Celio, C. I., Durlak, J., & Dymnicki, A. (2011). A meta-analysis of the impact of service-learning on students. *Journal of Experiential Education, 34,* 164–181. http://dx.doi.org/10.1177/105382591103400205

*Chang, M. J. (1996). *Racial diversity in higher education: Does a racially mixed student population affect educational outcomes?* (Unpublished doctoral dissertation). University of California, Los Angeles, Los Angeles, CA.

*Chang, M. J. (2002). The impact of an undergraduate diversity course requirement on students' racial views and attitudes. *The Journal of General Education, 51,* 21–42. http://dx.doi.org/10.1353/jge.2002.0002

*Chappell, C. (2014). Evaluating the effectiveness of an undergraduate multicultural course. *Psychology Learning & Teaching, 13,* 243–249. http://dx.doi.org/10.2304/plat.2014.13.3.243

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York, NY: Routledge. http://dx.doi.org/10.4324/9780203774441

Cole, D., & Zhou, J. (2014). Do diversity experiences help college students become more civically minded? Applying Banks' multicultural education framework. *Innovative Higher Education, 39,* 109–121. http://dx.doi.org/10.1007/s10755-013-9268-x

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 and Ethnic Minority Psychology, 17,* 397–405. http://dx.doi.org/10.1037/a0025433

Conway, J. M., Amel, E. L., & Gerwien, D. P. (2009). Teaching and learning in the social context: A meta-analysis of service-learning's effects on academic, personal, social, and citizenship outcomes. *Teaching of Psychology, 36,* 233–245. http://dx.doi.org/10.1080/00986280903172969

Crisp, R. J., & Turner, R. N. (2011). Cognitive adaptation to the experience of social and cultural diversity. *Psychological Bulletin, 137,* 242–266. http://dx.doi.org/10.1037/a0021840

Davies, K., Tropp, L. R., Aron, A., Pettigrew, T. F., & Wright, S. C. (2011). Cross-group friendships and intergroup attitudes: A meta-analytic review. *Personality and Social Psychology Review, 15,* 332–351. http://dx.doi.org/10.1177/1088868311411103

de Novais, J., & Spencer, G. (2019). Learning race to unlearn racism: The effects of ethnic-studies course-taking. *The Journal of Higher Education, 90,* 860–883.

Denson, N. (2009). Do curricular and cocurricular diversity activities influence racial bias? A meta-analysis. *Review of Educational Research, 79,* 805–838. http://dx.doi.org/10.3102/0034654309331551

Denson, N., & Bowman, N. A. (2017). Do diversity courses make a difference? A critical examination of college diversity coursework and student outcomes. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 35–84). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-48983-4_2

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

Engberg, M. E. (2004). Improving intergroup relations in higher education: A critical examination of the influence of educational interventions on racial bias. *Review of Educational Research, 74,* 473–524. http://dx.doi.org/10.3102/00346543074004473

Engberg, M. E. (2007). Educating the workforce for the 21st century: A cross-disciplinary analysis of the impact of the undergraduate experience on students' development of a pluralistic orientation. *Research in Higher Education, 48,* 283–317. http://dx.doi.org/10.1007/s11162-006-9027-2

Engberg, M. E., & Hurtado, S. (2011). Developing pluralistic skills and dispositions in college: Examining racial/ethnic group differences. *The Journal of Higher Education, 82,* 416–443. http://dx.doi.org/10.1353/jhe.2011.0025

Engberg, M. E., Hurtado, S., & Smith, G. C. (2007). Developing attitudes of acceptance toward lesbian, gay, and bisexual peers: Enlightenment, contact, and the college experience. *Journal of Gay & Lesbian Issues in Education, 4,* 49–77. http://dx.doi.org/10.1300/J367v04n03_05

Engberg, M. E., & Mayhew, M. J. (2007). The influence of first-year "success" courses on student learning and democratic outcomes. *Journal of College Student Development, 48,* 241–258. http://dx.doi.org/10.1353/csd.2007.0023

Gurin, P., Dey, E., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educational Review, 72,* 330–367. http://dx.doi.org/10.17763/haer.72.3.01151786u134n051

Gurin, P., Nagda, B. R. A., & Lopez, G. E. (2004). The benefits of diversity in education for democratic citizenship. *Journal of Social Issues, 60,* 17–34. http://dx.doi.org/10.1111/j.0022-4537.2004.00097.x

Gurin, P., Nagda, B. R. A., & Zúñiga, X. (2013). *Dialogue across difference: Practice, theory, and research on intergroup dialogue*. New York, NY: Russell Sage Foundation.

Harper, C. E., & Yeung, F. (2013). Perceptions of institutional commitment to diversity as a predictor of college students' openness to diverse perspectives. *Review of Higher Education: Journal of the Association for the Study of Higher Education, 37,* 25–44. http://dx.doi.org/10.1353/rhe.2013.0065

Hart Research Associates. (2016). *Recent trends in general education design, learning outcomes, and teaching approaches: Key findings from a survey among administrators at AAC&U member institutions*. Retrieved from the Association of American Colleges and Universities website http://www.aacu.org/about/2015-membersurvey

Hedges, L., & Olkin, I. (1985). *Statistical models for meta-analysis*. New York, NY: Academic Press.

Henderson-King, D., & Kaleta, A. (2000). Learning about social diversity: The undergraduate experience and intergroup tolerance. *The Journal of Higher Education, 71,* 142–164. http://dx.doi.org/10.1080/00221546.2000.11778832

Herzog, S. (2010). Gauging the effect of compositional and curricular diversity on freshman success. *New Directions for Institutional Research, 2010,* 73–97. http://dx.doi.org/10.1002/ir.323

Herzog, S., & Bowman, N. A. (Eds.). (2011). *Validity and limitations of college student self-report data. New directions for institutional research* (Vol. 110). Hoboken, NJ: Wiley.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557

Hurtado, S. (1994). The institutional climate for talented Latino students. *Research in Higher Education, 35,* 21–41. http://dx.doi.org/10.1007/BF02496660

Hurtado, S. (2005). The next generation of diversity and intergroup relations research. *Journal of Social Issues, 61,* 595–610. http://dx.doi.org/10.1111/j.1540-4560.2005.00422.x

Hurtado, S., Alvarez, C. L., Guillermo-Wann, C., Cuellar, M., & Arellano, L. (2012). A model for diverse learning environments. In J. C. Smart & M. B. Paulsen (Eds.), *Higher education: Handbook of theory and research* (pp. 41–122). New York, NY: Springer. http://dx.doi.org/10.1007/978-94-007-2950-6_2

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

Hyun, M. (1994, November). *Helping to promote racial understanding: Does it matter if you're Black or White*. Paper presented at the annual meeting of Association of the Study of Higher Education, Tucson, AZ.

Inkelas, K. K. (2004). Does participation in ethnic cocurricular activities facilitate a sense of ethnic awareness and understanding? A study of Asian Pacific American undergraduates. *Journal of College Student Development, 45,* 285–302. http://dx.doi.org/10.1353/csd.2004.0037

Jaschik, S. (2017, August 14). White nationalists rally at University of Virginia. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2017/08/14/white-nationalists-rally-university-virginia

Jayakumar, U. (2008). Can higher education meet the needs of an increasingly diverse and global society? Campus diversity and cross-cultural workforce competencies. *Harvard Educational Review, 78,* 615–651. http://dx.doi.org/10.17763/haer.78.4.b60031p350276699

Johnson, S. M., & Lollar, X. L. (2002). Diversity policy in higher education: The impact of college students' exposure to diversity on cultural awareness and political participation. *Journal of Education Policy, 17,* 305–320. http://dx.doi.org/10.1080/02680930210127577

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.

Kendall Brown, M. E. (2008). *A mixed methods examination of college students' intercultural development* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor, MI.

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

Lindsay, N. K. (2007). *Enhancing perpetual learning: The nexus between a liberal arts education and the disposition toward lifelong learning* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor, MI.

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

Loes, C., Pascarella, E., & Umbach, P. (2012). Effects of diversity experiences on critical thinking skills: Who benefits? *The Journal of Higher Education, 83,* 1–25. http://dx.doi.org/10.1353/jhe.2012.0001

Loes, C. N., Salisbury, M. H., & Pascarella, E. T. (2013). Diversity experiences and attitudes toward literacy: Is there a link? *The Journal of Higher Education, 84,* 834–865. http://dx.doi.org/10.1353/jhe.2013.0039

Lopez, G. E. (2004). Interethnic contact, curriculum, and attitudes in the first year of college. *Journal of Social Issues, 60,* 75–94. http://dx.doi.org/10.1111/j.0022-4537.2004.00100.x

Lopez, G. E., Gurin, P., & Nagda, B. A. (1998). Education and understanding structural causes for group inequalities. *Political Psychology, 19,* 305–329. http://dx.doi.org/10.1111/0162-895X.00106

MacPhee, D., Kreutzer, 7. C., & Fritz, 7. 7. (1994). Infusing a diversity perspective into human development courses. *Child Development, 65,* 699–715. http://dx.doi.org/10.2307/1131411

Martin, M. M. (2006). *The cognitive and affective outcomes of a cultural diversity in business course in higher education* (Unpublished doctoral dissertation) University of Missouri, Columbia, MO. http://dx.doi.org/10.32469/10355/4414

Mayhew, M. 7., & Engberg, M. E. (2010). Diversity and moral reasoning: How negative diverse peer interactions affect the development of moral reasoning in undergraduate students. *The Journal of Higher Education, 81,* 459–488. http://dx.doi.org/10.1080/00221546.2010.11779061

Mayhew, M. 7., Grunwald, H. E., & Dey, E. L. (2005). Curriculum matters: Creating a positive climate for diversity from the student perspective. *Research in Higher Education, 46,* 389–412. http://dx.doi.org/10.1007/s11162-005-2967-0

Mayhew, M. 7., Rockenbach, A. N., Bowman, N. A., Seifert, T. A., Wolniak, G. C., Pascarella, E. T., & Terenzini, P. Y. (2016). *How college affects students: 21st century evidence that higher education works* (Vol. *3*). San Francisco, CA: 7ossey-Bass.

Mayhew, M. 7., Seifert, T. A., & Pascarella, E. T. (2012). How the first year of college influences moral reasoning development for students in moral consolidation and moral transition. *Journal of College Student Development, 53,* 19–40. http://dx.doi.org/10.1353/csd.2012.0004

Mayhew, M. 7., Wolniak, G. C., & Pascarella, E. T. (2008). How educational practices affect the development of life-long learning orientations in traditionally-aged undergraduate students. *Research in Higher Education, 49,* 337–356. http://dx.doi.org/10.1007/s11162-007-9081-4

Milem, 7. E. (1994). College, students, and racial understanding. *Thought & Action, 9,* 51–92.

Milem, 7. F., Umbach, P. D., & Liang, C. T. (2004). Exploring the perpetuation hypothesis: The role of colleges and universities in desegregating society. *Journal of College Student Development, 45,* 688 –700. http://dx.doi.org/10.1353/csd.2004.0070

Nagda, B. R. A., Gurin, P., Sorensen, N., Gurin-Sands, C., & Osuna, S. M. (2009). From separate corners to dialogue and action. *Race and Social Problems, 1,* 45–55. http://dx.doi.org/10.1007/s12552-009-9002-6

National Survey of Student Engagement. (2011). *2019 topical module: Inclusiveness and engagement with cultural diversity*. Bloomington: Indiana University Center for Postsecondary Research.

Nelson, M. (2010). *Assessing the impact of diversity courses on students' values, attitudes and beliefs* (Unpublished doctoral dissertation). University of Southern California, Los Angeles, CA.

Nelson Laird, T. F. (2003). *Exploring "diversity" courses and their effects on aspects of students' identities* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor, MI.

Nelson Laird, T. F. (2005). College students' experiences with diversity and their effects on academic self-confidence, social agency, and disposition toward critical thinking. *Research in Higher Education, 46,* 365–387. http://dx.doi.org/10.1007/s11162-005-2966-1

Nelson Laird, T. F., Engberg, M. E., & Hurtado, S. (2005). Modeling accentuation effects: Enrolling in a diversity course and the importance of social action engagement. *The Journal of Higher Education, 76,* 448–476. http://dx.doi.org/10.1353/jhe.2005.0028

Nuñez, A. M. (2009). Latino students' transitions to college: A social and intercultural capital perspective. *Harvard Educational Review, 79,* 22–48. http://dx.doi.org/10.17763/haer.79.1.wh7164658k33w477

Orfield, G., Ee, 7., Frankenberg, E., & Siegel-Hawley, G. (2016). *Brown and 62: School segregation by race, poverty and state*. Los Angeles: The Civil Rights Project/Proyecto Derechos Civiles at UCLA.

Park, 7. 7. (2009). Are we satisfied? A look at student satisfaction with diversity at traditionally White institutions. *Review of Higher Education: Journal of the Association for the Study of Higher Education, 32,* 291–320. http://dx.doi.org/10.1353/rhe.0.0071

Pascarella, E. T., Palmer, B., Moye, M., & Pierson, C. T. (2001). Do diversity experiences influence the development of critical thinking? *Journal of College Student Development, 42,* 257–271.

Pascarella, E. T., Salisbury, M. H., & Blaich, C. (2013). Design and analysis in college impact research: Which counts more? *Journal of College Student Development, 54,* 329–335. http://dx.doi.org/10.1353/csd.2013.0032

Pascarella, E. T., Salisbury, M. H., Martin, G. L., & Blaich, C. (2012). Some complexities in the effects of diversity experiences on orientation toward social/political activism and political views in the first year of college. *The Journal of Higher Education, 83,* 467–496. http://dx.doi.org/10.1353/jhe.2012.0026

Pearson, M. A. (2012). *The impact of diversity courses on students' critical thinking skills* (Unpublished doctoral dissertation). University of Southern California, Los Angeles, CA.

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology, 90,* 175–181. http://dx.doi.org/10.1037/0021-9010.90.1.175

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. (2011). *When groups meet: The dynamics of intergroup contact*. New York, NY: Psychology Press.

Polanin, 7. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research, 86,* 207–236. http://dx.doi.org/10.3102/0034654315582067

Porter, S. R. (2013). Self-reported learning gains: A theory and test of college student survey response. *Research in Higher Education, 54,* 201–226. http://dx.doi.org/10.1007/s11162-012-9277-0

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

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

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.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). HLM 7 for Windows [Computer software]. Skokie, IL: Scientific Software International.

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

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638.

Saenz, V. B. (2005). *Breaking the cycle of segregation: Examining students' pre-college racial environments and their diversity experiences in college* (Unpublished doctoral dissertation). University of California, Los Angeles, Los Angeles, CA.

Saenz, V. B. (2010). Breaking the segregation cycle: Examining students' precollege racial environments and college diversity experiences. *Review of Higher Education: Journal of the Association for the Study of Higher Education, 34,* 1–37. http://dx.doi.org/10.1353/rhe.2010.0000

Saenz, V. B., Ngai, H. N., & Hurtado, S. (2007). Factors influencing positive interactions across race for African American, Asian American, Latino, and White college students. *Research in Higher Education, 48,* 1–38. http://dx.doi.org/10.1007/s11162-006-9026-3

Sciame-Giesecke, S., Roden, D., & Parkison, K. (2009). Infusing diversity into the curriculum: What are faculty members actually doing? *Journal ofDiversity in Higher Education, 2,* 156–165. http://dx.doi.org/10.1037/a0016042

Seifert, T. A., Bowman, N. A., Wolniak, G. C., Rockenbach, A. N., & Mayhew, M. J. (2017). Ten challenges and recommendations for advancing research on the effects of college on students. *AERA Open, 3,* 1–12. http://dx.doi.org/10.1177/2332858417701683

Simon, M., & Sidner, S. (2018, November 29). In 2008, there was hope. In 2018, there is hurt. This is America's state of hate. *CNN*. Retrieved from https://edition.cnn.com/2018/11/26/us/america-state-of-hate/index.html

Smith, H., Parr, R., Woods, R., Bauer, B., & Abraham, T. (2010). Five years after graduation: Undergraduate cross-group friendships and multicultural curriculum predict current attitudes and activities. *Journal of College Student Development, 51,* 385–402. http://dx.doi.org/10.1353/csd.0.0141

Stake, J. E., & Hoffmann, F. L. (2001). Changes in student social attitudes, activism, and personal confidence in higher education: The role of women's studies. *American Educational Research Journal, 38,* 411–436. http://dx.doi.org/10.3102/00028312038002411

Steiner, P. M., Cook, T. D., Shadish, W. R., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods, 15,* 250–267. http://dx.doi.org/10.1037/a0018719

Taylor, S. H. (1994). *Enhancing tolerance: The confluence of moral development with the college experience* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor, MI.

Tsui, L. (1999). Courses and instruction affecting critical thinking. *Research in Higher Education, 40,* 185–200. http://dx.doi.org/10.1023/A:1018734630124

VanHecke, J. R. (2006). *Responsible citizenship among college students: A study of predictive experiences and values* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor, MI.

van Laar, C., Sidanius, J., & Levin, S. (2008). Ethnic-related curricula and intergroup attitudes in college: Movement toward and away from the in-group. *Journal of Applied Social Psychology, 38,* 1601–1638. http://dx.doi.org/10.1111/j.1559-1816.2008.00361.x

Vogelgesang, L. J. (2001). *The Impact of college on the development of civic values: How do race and gender matter?* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Warchal, J. R. (1999). *An investigation of racial identity attitudes, racism, sexism, and homophobia in students enrolled in undergraduate diversity courses* (Unpublished doctoral dissertation). Lehigh University, Bethlehem, PA.

Warren, J. L. (2012). Does service-learning increase student learning? A meta-analysis. *Michigan Journal of Community Service Learning, 18,* 56–61.

Yeazel, M. R. (2008). *Relationship between critical thinking disposition and level of tolerance* (Unpublished doctoral dissertation). Capella University, Minneapolis, MN.

Zúñiga, X., Williams, E. A., & Berger, J. B. (2005). Action-oriented democratic outcomes: The impact of student involvement with campus diversity. *Journal of College Student Development, 46,* 660–678. http://dx.doi.org/10.1353/csd.2005.0069

Received August 14, 2019
Revision received March 30, 2020
Accepted April 7, 2020 ■

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.