Research Article

Psychological Science
2018, Vol. 29(4) 549–571
© The Author(s) 2018
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797617739704
www.psychologicalscience.org/PS

To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses

Victoria F. Sisk[1], Alexander P. Burgoyne[2], Jingze Sun[1], Jennifer L. Butler[1], and Brooke N. Macnamara[1]
[1]Department of Psychological Sciences, Case Western Reserve University, and
[2]Department of Psychology, Michigan State University

Abstract
Mind-sets (aka implicit theories) are beliefs about the nature of human attributes (e.g., intelligence). The theory holds that individuals with *growth* mind-sets (beliefs that attributes are malleable with effort) enjoy many positive outcomes—including higher academic achievement—while their peers who have *fixed* mind-sets experience negative outcomes. Given this relationship, interventions designed to increase students' growth mind-sets—thereby increasing their academic achievement—have been implemented in schools around the world. In our first meta-analysis (k = 273, N = 365,915), we examined the strength of the relationship between mind-set and academic achievement and potential moderating factors. In our second meta-analysis (k = 43, N = 57,155), we examined the effectiveness of mind-set interventions on academic achievement and potential moderating factors. Overall effects were weak for both meta-analyses. However, some results supported specific tenets of the theory, namely, that students with low socioeconomic status or who are academically at risk might benefit from mind-set interventions.

Keywords
mind-set, implicit theories, education, academic achievement, open data

Received 2/21/17; Revision accepted 10/9/17

According to *mind-set theory* (aka implicit theories; Dweck, 2006; Dweck, Chiu, & Hong, 1995), individuals vary in their beliefs about whether human attributes (e.g., intelligence) are stable or malleable. Individuals who believe attributes are stable have *fixed mind-sets* (aka entity theories), whereas those who believe attributes are malleable have *growth mind-sets* (aka incremental theories). According to mind-set theory, holding a fixed mind-set is detrimental for a variety of real-world outcomes, whereas holding a growth mind-set leads to a variety of positive outcomes, including weight loss (Burnette & Finkel, 2012), reaching international acclaim (Dweck, 2006), and achieving peace in the Middle East (Dweck, 2012, 2016).

Most frequently, mind-sets are researched in educational contexts (Burnette, O'Boyle, VanEpps, Pollack, & Finkel, 2013). Mind-set theory suggests that students with higher growth mind-sets have more adaptive psychological traits and behaviors (e.g., positive response to failure), which lead to greater academic achievement (e.g., Dweck, 2000). The theory also suggests that interventions designed to increase students' growth mind-sets will lead to greater academic achievement because there is a "powerful impact of growth mindset messages upon students' attainment" (Boaler, 2013, p. 143). These ideas have led to the establishment of nonprofit organizations (e.g., Project for Education Research That Scales [PERTS]), for-profit entities (e.g., Mindset Works, Inc.), schools purchasing mind-set intervention programs

Corresponding Author:
Brooke N. Macnamara, Department of Psychological Sciences, Case Western Reserve University, 11220 Bellflower Rd., Cleveland, OH 44106
E-mail: brooke.macnamara@case.edu

(e.g., Brainology), and millions of dollars in funding to individual researchers, nonprofit organizations, and for-profit companies (e.g., Bill and Melinda Gates Foundation,[1] Department of Education,[2] Institute of Educational Sciences[3]).

Given mind-set theory's impact on education, we sought to ask the following questions:

1. What is the magnitude of the relationship between mind-sets and academic achievement, and under which circumstances does the relationship strengthen or weaken?
2. Do mind-set interventions positively impact academic achievement, and under which circumstances does the impact increase or decrease?

To answer these questions, we conducted two meta-analyses to (a) estimate the sizes of these effects and whether they are consistent across studies, (b) examine potential moderating factors, and (c) empirically evaluate the theory.

Meta-Analysis 1: The Relationship Between Mind-Sets and Academic Achievement

Mind-set theory suggests that mind-sets play critical roles in academic achievement (Rattan, Savani, Chugh, & Dweck, 2015). For example, Dweck (2008) stated, "what students believe about their brains — whether they see their intelligence as something that's fixed or something that can grow and change — has profound effects on their motivation, learning, and school achievement" (para. 2). In the first meta-analysis, we examined the magnitude of the relationship between mind-sets and academic achievement.

Next, we investigated potential moderators. We examined *academic risk status* because the theory holds that having a growth mind-set is especially important for at-risk students (e.g., Paunesku et al., 2015) and students facing situational challenges such as school transitions (e.g., Yeager & Dweck, 2012). According to the theory, students with growth mind-sets will interpret struggles as learning opportunities, while students with fixed mind-sets will be "devastated by setbacks" (Dweck, 2008, para. 2; see also Burnette et al., 2013). Similarly, although the theory is not linked to a particular age, some researchers suggest that mind-sets are particularly influential during the tumultuous period of adolescence when students face new challenges (Blackwell, Trzesniewski, & Dweck, 2007). To assess the importance of this moderator, we examined *student developmental stage*. Additionally, we examined *socioeconomic status* (SES) because some research (e.g., Claro, Paunesku, & Dweck, 2016) has suggested that holding a growth mind-set is especially beneficial for low-SES students' academic success.

We examined the *type of academic achievement measure* because the effect might differ, for example, between course grades and standardized tests. Additionally, we investigated the possibility that if students with growth mind-sets are taking more challenging courses (see Romero, Master, Paunesku, Dweck, & Gross, 2014), then the relationship could be suppressed when the measure of achievement also reflects students' course selection.

Finally, we tested whether publication bias is problematic within the mind-set-in-education literature. Publication bias occurs when some results are systematically less likely to be published than others (e.g., studies that find small or null effects; Rosenthal, 1979).

Method

We designed the meta-analysis and report the results in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009).

Inclusion criteria, literature search, and coding. We searched for studies for both meta-analyses in a single search. The criteria for including a study in Meta-Analysis 1 were as follows:

- A measure of a belief about one or more human attributes (e.g., intelligence) as fixed or malleable—henceforth *mind-set*—was collected.
- A mind-set measure was collected prior to or without a mind-set intervention.
- A measure of academic achievement—course exam (e.g., midterm exam), course grade, average of course grades (e.g., grade point average, or GPA), or standardized test performance—was collected prior to or without a mind-set intervention.
- A bivariate correlation coefficient reflecting the relationship between mind-set and academic achievement was reported, or enough information was provided to compute this effect size.
- The methods and results were in English.

Mind-set is typically measured using participants' responses to statements such as, "No matter who you are, you can significantly change your intelligence level" and "You have a certain amount of intelligence, and you can't really do much to change it" (reverse scored) using a Likert scale (e.g., Dweck, 2006). The more students agree with statements about the malleability of an attribute, the more of a growth mind-set they hold. Measures of beliefs about the importance of effort without corresponding beliefs about the malleability of one or more human attributes were not included. Likewise, mind-set of willpower was not included because (a) willpower refers to exerted control rather than an attribute, and (b) mind-set of willpower focuses on beliefs about whether willpower is limited or not limited rather than whether an attribute is stable or changes with effort.

To identify studies meeting these criteria and the criteria set forth for Meta-Analysis 2, we systematically searched for relevant published and unpublished articles in psychology, education, and other disciplines through October 28, 2016 (for a flowchart designed according to the PRISMA specifications, see Fig. 1). We also e-mailed authors of articles on mind-set (N = 137) and asked that they forward the e-mail to colleagues who might have conducted relevant studies. Further, we contacted organizations dedicated to intervention-in-education research (e.g., PERTS) to request information relevant to our meta-analysis that was not accessible (e.g., unpublished data), and we posted requests for unpublished data on a Society for Personality and Social Psychology forum. We accepted new data from these calls through January 11, 2017. Following our search stop date, we evaluated studies for eligibility and coded each study and the measures collected in it for reference information, student characteristics, methodological characteristics, and results (the data file is available at osf.io/453ds). We included updates to our existing records until analyses began on February 1, 2017.

Our search included 15,867 novel records. After examining these records and discarding obviously irrelevant ones (e.g., literature reviews, commentaries), we identified 129 studies that met all the inclusion criteria for Meta-Analysis 1. These studies included 162 independent samples, with 273 effect sizes and a total sample size of 365,915 students. In cases where authors reported effects associated with multiple measures of mind-set (e.g., a fixed mind-set scale and a growth mind-set scale) or multiple measures of academic achievement (e.g., GPA and performance on a standardized test), we adjusted for dependent samples by using a method based on that of Cheung and Chan (2004, 2008). This method statistically adjusts (lowers) the associated sample size because of dependent effects being partially redundant, which reduces the weight of these effect sizes in the meta-analysis so as not to overly contribute to the model. For a list of studies included in the meta-analysis, see the Supplemental Material available online or the file at osf.io/453ds. For additional characteristics of Meta-Analysis 1, see Table 1.

Effect sizes. To measure the magnitude of the relationship, we used the correlation as the measure of effect size. For most studies, the authors reported a Pearson's correlation coefficient; for studies in which the authors reported group-level comparisons (e.g., students holding a growth mind-set vs. a fixed mind-set), we converted standardized mean differences (Cohen's ds) to biserial correlations (r_bs; Becker, 1986; Hunter & Schmidt, 1990). There was not a significant difference in effect sizes between studies that reported group-level comparisons and those that used continuous variables, p = .463. Most studies' authors coded higher scores on the mind-set measure as reflecting more of a growth mind-set. When authors used a mind-set measure where higher scores reflected more of a fixed mind-set, we reversed the sign of the correlation before analyzing the data. We also reversed the sign of the correlation in the rare cases where lower scores on a measure of academic achievement reflected better performance. For instance, in Germany, lower grades reflect better performance. Thus, all effect sizes were coded such that a positive correlation reflected a positive relationship between growth mind-set and academic achievement.

Moderator variables

Developmental stage. There were three levels of developmental stage: children (primary school students), adolescents (middle school, junior high school, and high school students), and adults (e.g., postsecondary students). Studies that included students in multiple categories (e.g., students in both primary school and junior high school in a single sample) were not included in this moderator analysis.

Academic risk status. There were three levels of academic risk status: high (at risk of failing; e.g., students who previously failed courses), moderate (facing a situational challenge; e.g., transitioning to a new school, a member of a stereotyped group under a stereotype threat manipulation), and low (no indicators that students were at risk). Each sample was categorized on the basis of the majority (> 50%) of the students in the sample. If we could obtain separate effect sizes for each subsample in a study based on risk level (e.g., an effect was available for the high-risk students as well as the remaining low-risk students), we did so and entered those effects as independent samples. If effects were available only for the entire sample and a high-risk subgroup, we replaced the entire sample with the high-risk subgroup when examining this moderator.

We did not code minority students or female students as academically at-risk samples unless they were under a relevant stereotype threat manipulation. While students can experience stereotype threat in natural

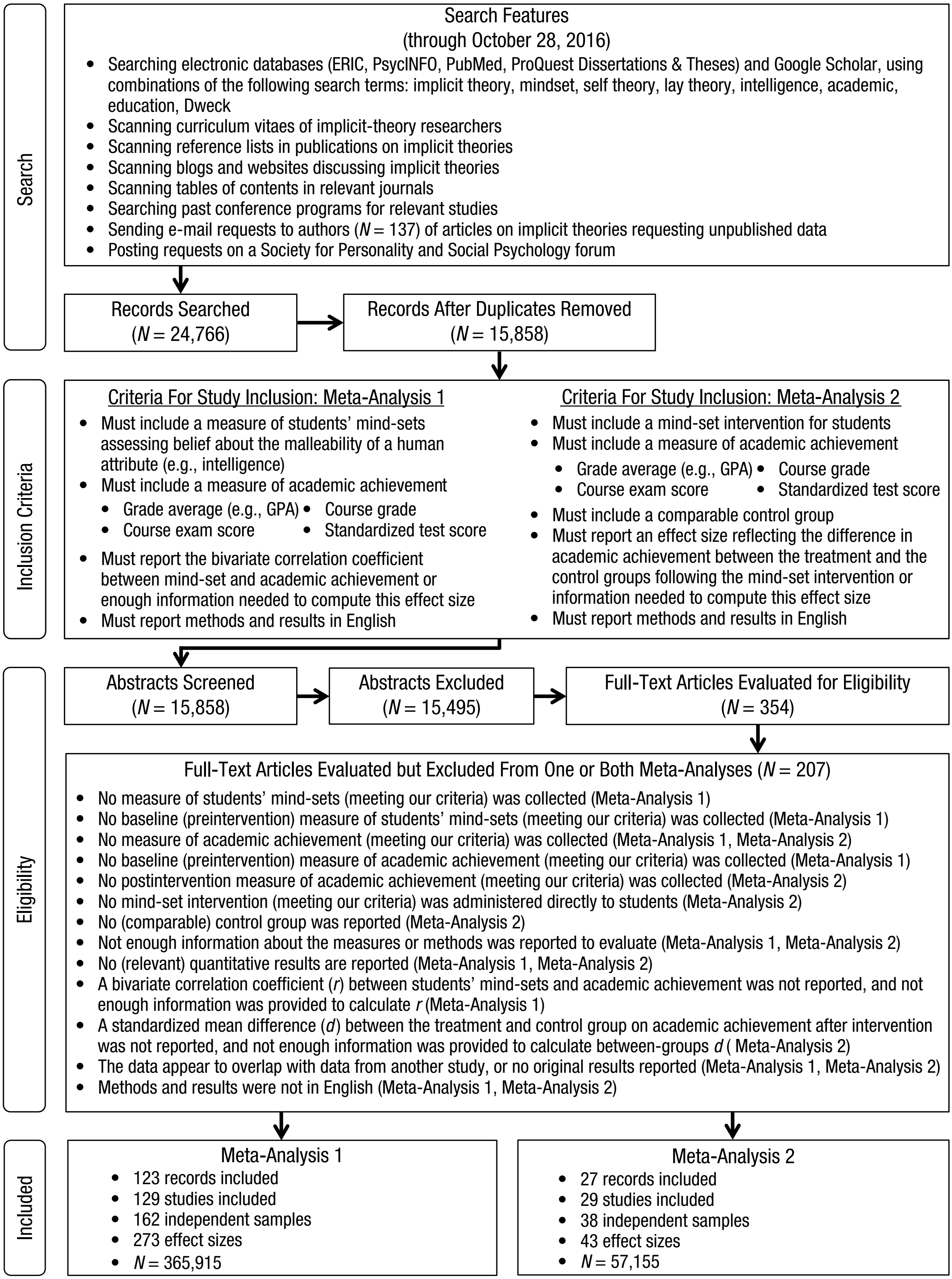

Fig. 1. Flow chart of the literature search and study coding. Three articles were included in both Meta-Analysis 1 and Meta-Analysis 2 (i.e., 147 unique records are included in the two meta-analyses). GPA = grade point average.

Table 1. Descriptive Characteristics for Meta-Analysis 1

Study characteristic	Number of effect sizes (k = 273)	Number of participants (N = 365,915)
Developmental stage[a]		
Children	50	8,118
Adolescents	126	332,240
Adults	89	21,673
Academic risk status[b]		
Low	208	346,043
Moderate	55	19,215
High	6	218
Socioeconomic status[b]		
Low	33	173,614
Not low	62	27,160
Mind-set type		
Intelligence	167	335,560
Other attribute (e.g., math ability)	106	30,355
Academic achievement measure		
Course exam	15	9,318
Course grade	51	11,384
Average grades (i.e., grade point average)	82	46,986
Standardized test	125	298,227
Laboratory measures	24	2,121
Publication status		
Published	116	323,040
Unpublished	157	42,875

[a]For this characteristic, some effect sizes are excluded because they included a wide age range.
[b]For this characteristic, one or more effect sizes were excluded because this information was not provided in the study.

environments, we cannot know whether this occurred in each study without stereotype threat being measured or manipulated. Some mind-set researchers have categorized minority students or female students as academically at risk and analyzed their results separately without measuring or manipulating stereotype threat. However, this categorization was not suitable in the present meta-analysis for multiple reasons. First, the cultural context varies across studies (Meta-Analysis 1, for instance, includes studies from over 20 different countries; see the Supplemental Material for more detail). Cultural conceptualizations of ethnic minorities vary across countries, such that identifying and categorizing at-risk groups in each sample to compare across studies would not be feasible or precise. Second, the educational contexts vary across studies in such a way that the association between academic risk and minority status is likely not consistent. In other words, the type and degree of challenge that minority students face depends on their educational environment, which varied across studies. Third, levels of achievement vary across studies such that academic risk for minority students and women is not constant. That is, low-achieving minority students do not have the same level of academic risk as high-achieving minority students. Fourth, minority status can be confounded with SES, which we coded separately. The effect of student ethnicity and gender on the relationship between growth mind-set and academic achievement is an important research question. However, we did not have the level of detail that would allow us to conduct a meaningful moderator analysis on ethnicity or gender as risk factors.

Socioeconomic status. There were two levels of SES: low SES (e.g., students qualified for reduced-price lunch) and not low SES (i.e., middle class or higher). Each study was categorized on the basis of the majority (> 50%) of the students in the sample. Studies not reporting student-level SES were not included in the SES moderator analysis.

Type of academic achievement measure. There were four levels of academic achievement measure: standardized test (e.g., Iowa Test of Basic Skills, SAT), and three pertaining to course performance—course exam (e.g., final exam score), course grade (e.g., math course grade), and cumulative or current GPA. When studies included multiple standardized test scores (e.g., verbal SAT, quantitative SAT, total SAT), we used the combined score when available. When studies included multiple course performance measures, we used the measure that provided the most comprehensive measure of academic achievement. That is, we used GPA when available because this provides the most information about a students' course performance. Likewise, we used course grades over course exams. We did not include enrollment status (full time vs. part time) or number of absences as academic achievement measures because they are not readily comparable with course performance or standardized test performance.

Some studies included measures of academic achievement administered by a researcher (e.g., practice questions on the GRE, a researcher-designed course-relevant test) as a proxy for academic achievement. We did not include researcher-designed tests as course exams if they were irrelevant to students' coursework (e.g., trivia quizzes, reading comprehension of the mind-set stimulus, worksheets on topics described as outside students' curricula). We present the results with and without laboratory measures because performance on these measures does not contribute to students' academic records.

Developmental stage as a moderator of mind-set on GPA. If students with a growth mind-set select more challenging courses or schools, it is possible that their GPAs would not be significantly higher than fixed mind-set students taking easier classes, leading to the relationship between mind-set and GPA being suppressed, especially for older students who have more opportunities for course selection. This suppression would affect only GPA, which reflects students' course selections. It would not suppress the effect on course exams or course grades because all students in the sample are in the same course. It would not suppress the effect on standardized tests because, if anything, students taking more challenging courses will be better prepared for standardized tests than students not exposed to higher-level material. We therefore also examine the interaction between mind-set and developmental stage on GPA.

Meta-analytic procedure. The meta-analysis involved four steps. The first step was to obtain correlations between mind-set of a human attribute and academic achievement, along with their sampling error variances. The second step was to search for extreme values. We defined outliers as effect sizes whose residuals had z scores of 3 or greater. There were no outliers for this meta-analysis. The third step was to use random-effects meta-analysis modeling, which assumes meaningful differences across studies; estimate the meta-analytic mean distribution of effects and heterogeneity in the effect sizes; and then test whether some of the heterogeneity was predictable from moderator variables, using mixed-effects meta-analysis modeling. The final step was to perform publication bias analyses.

We used the Comprehensive Meta Analysis Version 2 (Borenstein, Hedges, Higgins, & Rothstein, 2005) software package to conduct the meta-analyses. We used Comprehensive Meta Analysis, the R Project for Statistical Computing (www.r-project.org), the p-curve web application (http://www.p-curve.com/app), and the p-uniform web application (https://rvanaert.shinyapps.io/p-uniform) to conduct the publication bias analyses. (See also the Supplemental Material.)

Results

The model consists of 273 effect sizes. Effect sizes are weighted by the inverse of the variance, which includes both between-studies variance and within-study variances. Within-study variance accounts for the sample sizes, such that smaller studies are given less weight while larger studies are given more weight in the model. The majority of effect-size–associated adjusted Ns are ≥ 90. The mean adjusted N associated with this model's effect sizes is 1,429.

The meta-analytic average correlation (i.e., the average of various population effects) between growth mind-set and academic achievement is $\bar{r}$ = .10, 95% confidence interval (CI) = [.08, .13], p < .001. We did not correct individual effect sizes for the attenuating effect of measurement error (i.e., measurement unreliability), because very few studies in the meta-analysis reported a reliability estimate for mind-set. However, measures of mind-set have typically been found to have acceptable reliability greater than .80 (see, e.g., Dweck et al., 1995). If we assume reliability of .80, the meta-analytic average correlation between mind-set and academic achievement is $\bar{r}$ = .12, 95% CI = [.09, .14].

Figure 2 shows that 157 of the 273 effect sizes (58%) are not significantly different from zero. Another 16 effect sizes (6%) are significantly different from zero but negative, indicating that growth mind-sets were associated with worse academic achievement. The remaining 100 effect sizes (37%) are significantly different from zero and positive, indicating that growth mind-sets were positively associated with academic achievement. As can be seen in Figure 2, the effect sizes are not consistent across studies. The I^2 statistic specifies the percentage of the between-studies variability in effect sizes that is due to heterogeneity rather than random error. The I^2 statistic, I^2 = 96.29 (τ^2 = .025), demonstrated a very large proportion of heterogeneity in the effect sizes, indicating that the true effect of a given study could be substantially higher or lower than the meta-analytic average. We investigated the source of this heterogeneity through the moderator analyses reported next.

Moderator analyses

Student factors. The developmental stage of the students was a statistically significant moderator, $Q(2)$ = 72.84, p < .001. Eight effect sizes associated with samples with a wide age range encompassing more than one of the developmental levels (e.g., children and adolescents) were excluded from this analysis. The average correlation between mind-set and academic achievement was $\bar{r}$ = .19, 95% CI = [.16, .23], p < .001, for children; $\bar{r}$ = .15, 95% CI = [.12, .18], p < .001, for adolescents; and $\bar{r}$ = .02, 95% CI = [−.005, .05], p = .110, for adults. Post hoc follow-up analyses were conducted to determine the source of the difference using a corrected alpha of .05/3 = .017 for multiple comparisons. Adults differed significantly from both adolescents, $Q(1)$ = 39.89, p < .001, and children, $Q(1)$ = 58.37, p < .001. Adolescents and children did not differ significantly from each other, $Q(1)$ = 3.47, p = .063.

Academic risk status was not a significant moderator, $Q(2)$ = 0.22, p = .895. Four effect sizes did not have a sample description other than age and location and thus were removed from this analysis. The average correlation between mind-set and academic achievement was $\bar{r}$ = .11, 95% CI = [.08, .13], p < .001, for low-risk students; $\bar{r}$ = .11, 95% CI = [.07, .16], p < .001, for moderately at-risk students; and $\bar{r}$ = .08, 95% CI = [−.04, .21], p = .196, for highly at-risk students.

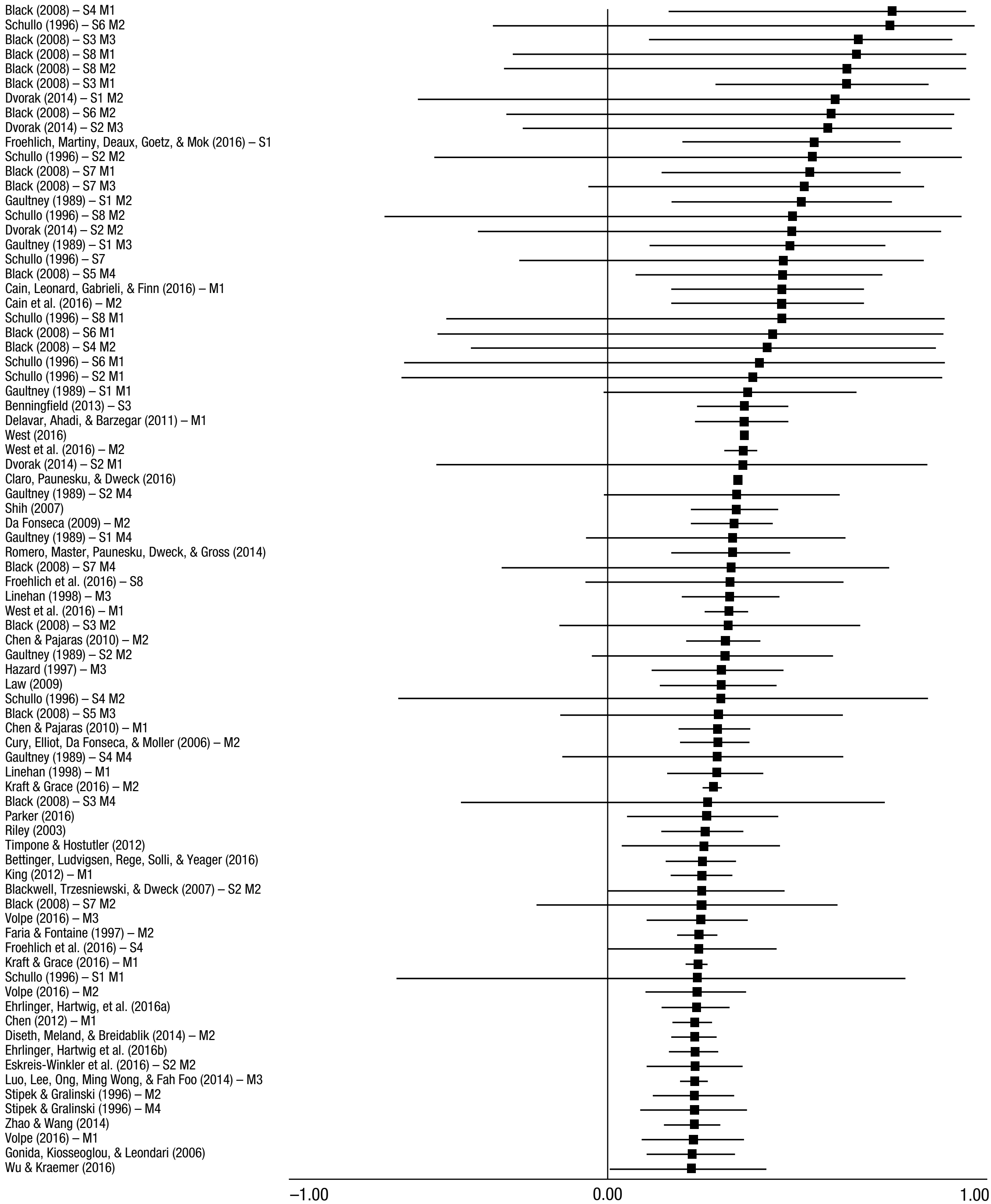

Fig. 2 *(continued on next page)*

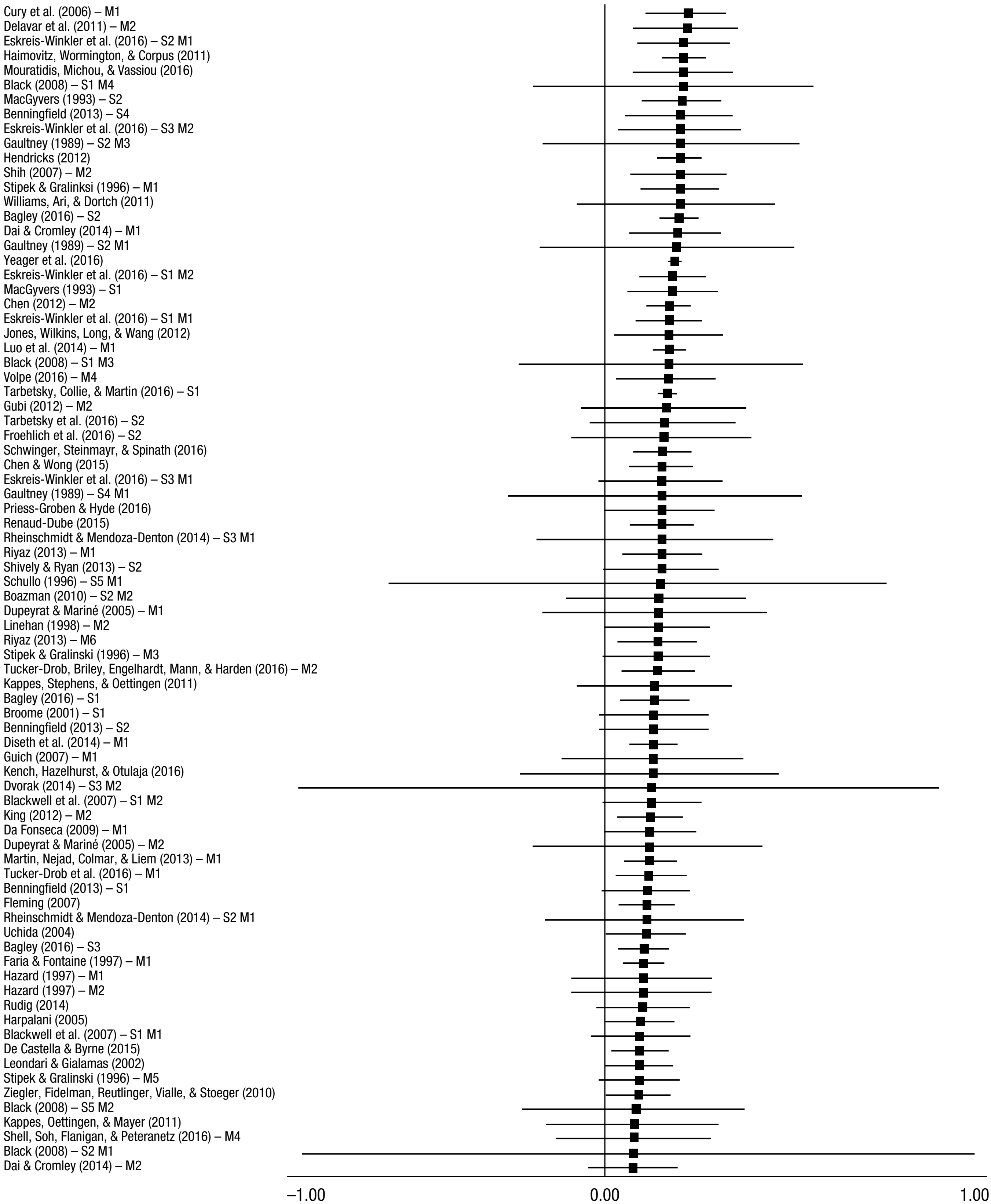

Fig. 2 *(continued on next page)*

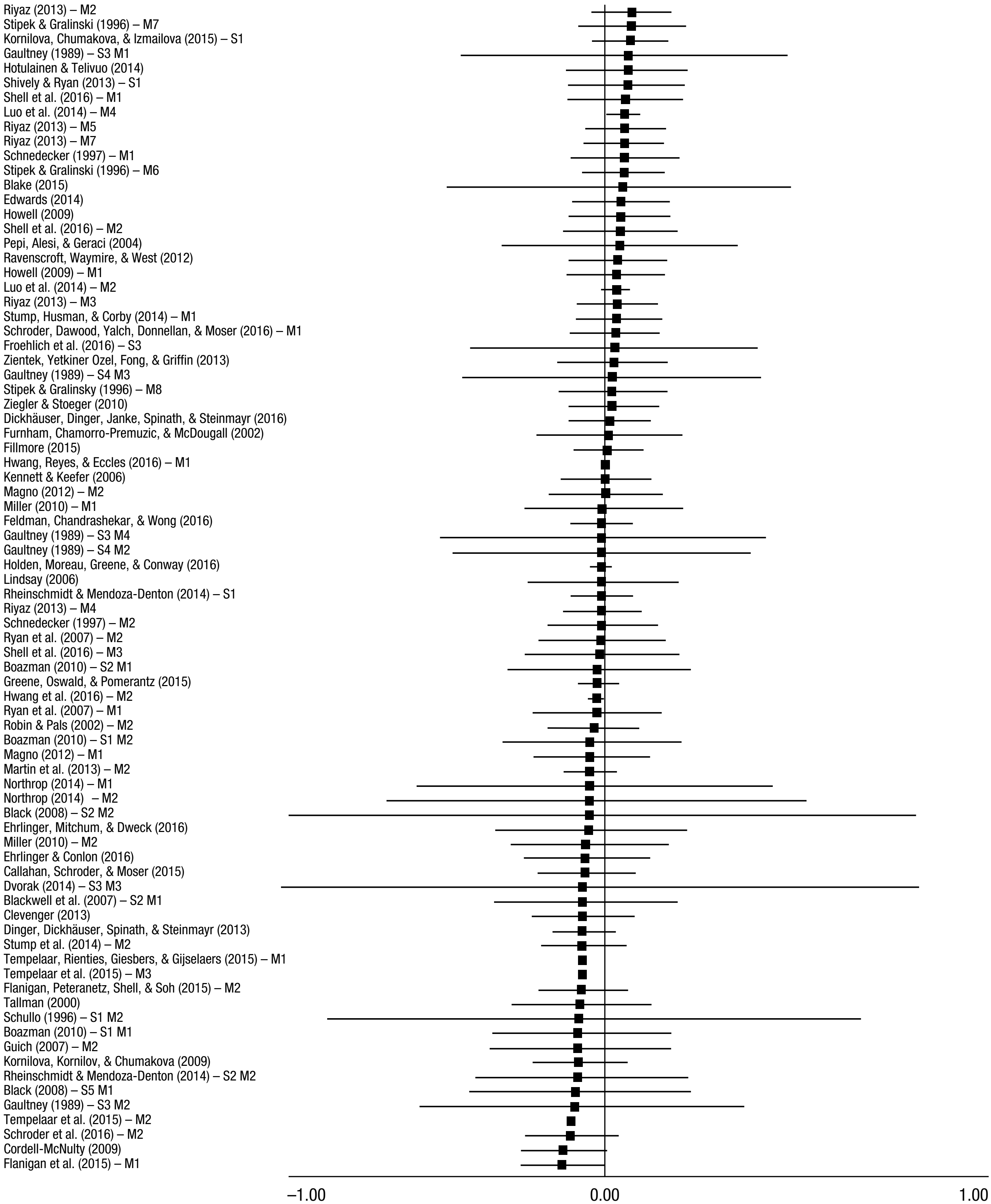

Fig. 2 *(continued on next page)*

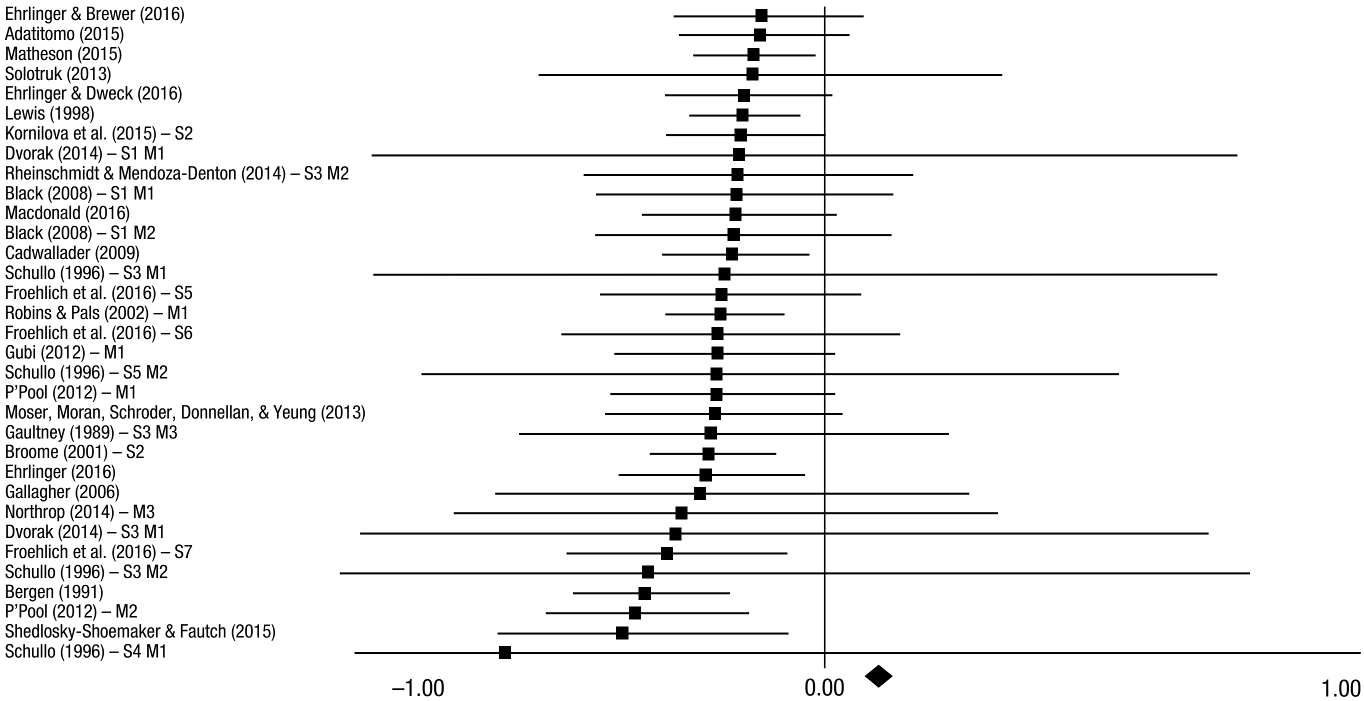

Fig. 2. Correlations between growth mind-set and academic achievement. Correlations (squares) and 95% confidence intervals (error bars) are displayed for all effects entered into Meta-Analysis 1. The diamond on the bottom row represents the meta-analytically weighted mean correlation coefficient. For studies with multiple independent samples, the result for each sample (S1, S2, etc.) is reported separately. Similarly, for studies with multiple measures, the result for each measure (M1, M2, etc.) is reported separately. Multiple measures were adjusted for dependency. See the Supplemental Material available online for details on all references.

Socioeconomic status was not a significant moderator, $Q(1) = 1.48$, $p = .223$. The 95 effect sizes associated with reported student-level SES were included in this analysis. The average correlation between mind-set and academic achievement was $\bar{r} = .17$, 95% CI = [.10, .23], $p < .001$, for low-SES students, and $\bar{r} = .12$, 95% CI = [.09, .16], $p < .001$, for middle-class and higher students.

Academic achievement measure. The measure of academic achievement used was not a statistically significant moderator, $Q(3) = 6.18$, $p = .103$. The average correlation between mind-set and academic achievement was $\bar{r} = .08$, 95% CI = [.01, .15], $p = .027$, for studies that used a course exam; $\bar{r} = .13$, 95% CI = [.09, .16], $p < .001$, for studies that used a course grade; $\bar{r} = .08$, 95% CI = [.05, .11], $p < .001$, for studies that used GPA; and $\bar{r} = .12$, 95% CI = [.09, .15], $p < .001$, for studies that used a standardized test.

Twenty-four effect sizes reflected the relationship between mind-set and a measure of academic achievement that was laboratory based. These included researcher-designed tests supposed to reflect comprehension of course-specific content (coded as a course exam) and standardized tests and portions of standardized tests administered by researchers in a laboratory setting (coded as standardized tests). Excluding the 24 effect sizes where the measure of academic achievement was a laboratory-based measure did not change the overall results. The overall meta-analytic average correlation with these effect sizes removed was $\bar{r} = .11$, 95% CI = [.08, .13], $p < .001$ (compare with $\bar{r} = .10$, 95% CI = [.08, .13], $p < .001$, when including these effect sizes). Excluding laboratory-based effect sizes did change the pattern of results for the academic achievement measure moderator. Without laboratory measures, the relationship between mind-set and academic achievement varied significantly on the basis of the academic achievement measure used, $Q(3) = 13.12$, $p = .004$. Specifically, removing researcher-designed course exams lowered the average correlation between mind-set and course exam performance, $\bar{r} = .04$, 95% CI = [−.02, .11], $p = .178$ (compare with when these effect sizes were included, $\bar{r} = .08$, 95% CI = [.01, .15], $p = .027$). The correlation between mind-set and academic achievement was similar for standardized tests regardless of whether laboratory-based effect sizes were excluded, $\bar{r} = .14$, 95% CI = [.11, .17], $p < .001$, or included, $\bar{r} = .12$, 95% CI = [.09, .15], $p < .001$. Post hoc follow-up tests using an adjusted alpha of .008 for multiple comparisons revealed that when laboratory-based measures were removed, course exams and GPA differed significantly from standardized tests, $Q(1) = 7.14$,

p = .0075, and $Q(1) = 8.36$, p = .004, respectively. No other pairwise comparisons were significant at the .008 level, all *p*s > .029.

Developmental stage as a moderator of mind-set on GPA. If students with growth mind-sets select more challenging courses, we would expect two patterns of results. First, the relationship between mind-set and academic achievement would not be suppressed for children who typically have little control over their course selection, somewhat suppressed for adolescents who have more course selection opportunities, and most suppressed for adults who have the most opportunities for course selection. The other pattern of results we would expect if students with growth mind-sets are selecting more challenging courses is that the relationship between mind-set and academic achievement will be suppressed when the measure of academic achievement is GPA, because GPA reflects performance in the students' selected courses. The relationship should not be suppressed for this reason when the measure is course grade or course exam because, in these cases, all students are taking the same courses. The relationship should also not be suppressed for this reason when the measure is standardized test performance because, if anything, students exposed to higher-level material should perform better than students taking less challenging courses.

Only four effect sizes were associated with children's mind-set and GPA and thus were excluded from this analysis as were samples associated with a wide range of ages (e.g., children and adolescents in the same sample). The difference between adolescents and adults was significant, $Q(1) = 17.01$, $p < .001$, with adolescents exhibiting a stronger relationship between mind-set and GPA, $\bar{r} = .12$, 95% CI = [.07, .16], $p < .001$, than adults, $\bar{r} = .002$, 95% CI = [−.03, .03], p = .892. This pattern of results supports the suppression hypothesis. However, the relationship between mind-set and GPA did not differ from the relationship between mind-set and other measures of academic achievement where course selection is unlikely to affect the relationship, $Q(3) = 6.18$, p = .103. Additionally, the relationship between mind-set and academic achievement was lowest for adults across all measures of academic achievement ($\bar{r} = .02$, 95% CI = [−.005, .05], p = .110), including those where course selection is unlikely to affect this relationship.

Publication bias analyses. Publication bias threatens the validity of published research by masking small and null effects. We conducted three types of publication bias analyses. The first analysis tested whether studies finding weak effect sizes were less likely to be published than studies finding stronger effects. We tested this via moderator analysis comparing effect sizes from published versus unpublished studies. The second analysis tested whether selective reporting of results (i.e., *p*-hacking) was responsible for significant effects in the published literature. We tested this via *p*-curve analysis. The third analysis tested whether our meta-analysis is overestimating the meta-analytic mean effect size due to missing unpublished data we were unable to obtain and include. We tested this via Egger's regression.

Moderator analysis. We tested whether published studies, on average, report larger effect sizes than studies that remain unpublished. Unpublished studies included manuscripts in preparation to submit, manuscripts submitted but not yet accepted, conference papers and posters, and studies and manuscripts that have remained unpublished. Studies may remain unpublished for multiple reasons. For example, a study may remain unpublished because its methodology is weak. That is, the author may choose to not submit the study for publication, or the study may be rejected for publication because of methodological shortcomings. Alternatively, studies may remain unpublished because they found small effects, null effects, or effects in the nonpredicted direction. That is, studies with results that do not strongly support a particular hypothesis may not be submitted for publication or may be rejected from publication. If this occurs, the overall effect size is overestimated within the published literature.

A moderator analysis revealed that the 157 correlations between mind-set and academic achievement from unpublished studies (median study sample size = 122) were not significantly different from the 116 correlations from published studies (median study sample size = 245), $Q(1) = 0.65$, p = .420. The average correlation between mind-set and academic achievement was $\bar{r} = .09$, 95% CI = [.07, .12], $p < .001$, for unpublished studies and $\bar{r} = .11$, 95% CI = [.08, .14], $p < .001$, for published studies.

p-*curve analysis.* We tested whether the source of published significant effects was due to *p*-hacking; that is, selective reporting of results (e.g., when authors conduct multiple analyses on the same data set but report only significant effects) or collecting data until a nonsignificant effect becomes significant (though see Bishop & Thompson, 2016).

The logic underlying the *p*-curve analysis is as follows. A *p*-curve depicts the distribution of statistically significant ($p < .05$) p values for a set of studies (Simonsohn, Nelson, & Simmons, 2014a). The *p*-curve analysis tests the skew of the *p*-value distribution. Studies demonstrating true effects will yield a right-skewed *p*-curve, indicating more lower significant p values (e.g., p = .001) than marginally significant p values (e.g., p = .049). Studies demonstrating null effects will yield a uniform (i.e., "flat") *p*-curve. Studies with predominately *p*-hacked effects will yield a left-skewed *p*-curve, indicating more marginally significant *p* values than lower significant *p* values (Simonsohn, Nelson, & Simmons, 2014a, 2014b). The results of this test should be interpreted with caution because *p*-curve estimates may be affected by large amounts of heterogeneity (van Aert, Wicherts, & van Assen, 2016).

We conducted the *p*-curve analyses in two steps. First, we classified effect sizes as either published or unpublished. If authors provided data that were not reported in a published study, the effect sizes calculated using those data were classified as unpublished for the purposes of this analysis. Simonsohn et al. (2014b) recommended that when choices among multiple effects must be made to adhere to a prespecified selection rule for the first (i.e., primary) and second (i.e., robustness) analyses (Simonsohn et al., 2014a). When there were multiple effect sizes from the same study included in the meta-analysis, we randomly selected one effect size from each study for the primary analysis and then randomly selected a different effect size for the robustness analysis. This process ensured that studies with a large number of effect sizes did not have undue influence on the analyses. Next, we ran primary and robustness *p*-curve analyses for the published effect sizes using the *p*-curve web application (http://www.p-curve.com/app).

The results of the primary *p*-curve analysis are presented in Figure 3. The result of the half *p*-curve test was $Z = -16.14$, $p < .0001$, and the result of the full *p*-curve test was $Z = -16.20$, $p < .0001$. These results suggest that the *p*-curve is significantly right-skewed, indicating evidential value. These results were corroborated by the robustness analysis (half: $Z = -21.16$, $p < .0001$; full: $Z = -20.25$, $p < .0001$). The primary and robustness *p*-curve analyses estimated that after correcting for selective reporting, the included studies had an estimated power of 99%. For additional figures and an index of the effect sizes that were entered into each analysis, see the Supplemental Material.

Egger's regression. We tested whether our meta-analysis was affected by missing studies. We found a considerable number of unpublished studies to include in the meta-analysis. However, during our search process, we became aware of unpublished studies that appeared to fit our inclusion criteria that we were unable to obtain. To examine how much these missing studies potentially affected our meta-analysis, we conducted Egger's regression (Egger, Smith, Schneider, & Minder, 1997). If a meta-analysis is unaffected by publication bias, larger studies' effect sizes will cluster around the mean effect size, while smaller studies' effect sizes (containing more sampling error) will be randomly dispersed around the mean (i.e., higher and lower than the mean). If a meta-analysis is affected by publication bias, smaller studies will contribute significantly more effect sizes with higher-than-average effects than smaller-than-average effects. Egger's regression was significant, $B_0 = -2.98$, 95% CI = [−3.58,

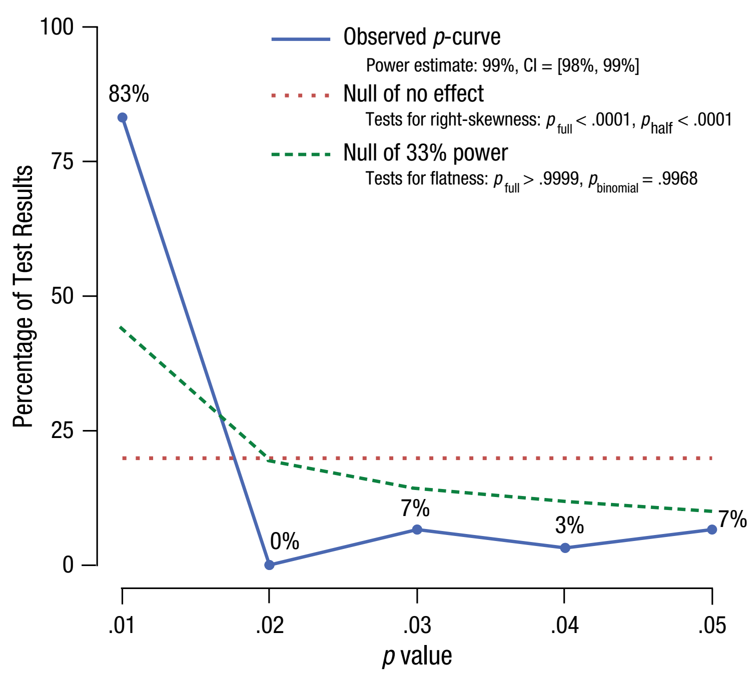

Fig. 3. Primary *p*-curve analysis for Meta-Analysis 1.

−2.39], $t(271) = 9.92$, $p < .001$, suggesting that the meta-analysis is likely overestimating the relationship between mind-set and academic achievement. However, Egger's regression is prone to Type I errors when heterogeneity is high (Sterne et al., 2011), as is the case in the current meta-analysis. Thus, this result should be interpreted with caution.

Discussion

The meta-analytic average correlation between growth mind-set and academic achievement was very weak—$\bar{r}$ = .10. This result is almost identical to the meta-analytic average correlation found between mind-set and achievement across achievement domains: $\bar{r}$ = .095 (Burnette et al., 2013). However, the overall effect is overshadowed by the high degree of heterogeneity.

Moderators were limited in accounting for this variance. Academic risk status and SES did not affect the relationship. Developmental stage moderated the relationship, though the effect remained weak for all subgroups and nonsignificant for adults. This pattern held when examining only GPA as the outcome, and GPA did not differ from other measures of academic achievement. Thus, there is limited evidence for a suppression effect due to students with growth mind-sets potentially selecting more challenging courses.

Growth-mind-set interventions in education are predicated on the relationship between mind-sets and academic achievement. However, it is possible that despite generally weak relationships between students' naturally held mind-sets and academic achievement, interventions promoting growth mind-sets might still be effective, especially for certain subgroups. We examined the effectiveness of growth-mind-set interventions on academic achievement next.

Meta-Analysis 2: The Effect of Growth-Mind-Set Interventions on Academic Achievement

Growth-mind-set interventions have been suggested as a way for students to earn higher grades and score higher on standardized tests (see mindsetscholarsnetwork.org/learning-mindsets/growth-mindset/). To examine the effectiveness of these interventions, we estimated the standardized mean differences in academic achievement between students who received a growth-mind-set intervention and students who did not.

To investigate potential moderators, we tested the same three student-related factors as in Meta-Analysis 1: developmental stage, academic risk status, and SES as well as control- and intervention-related methodological factors. We examined the type of control group (active control, passive control, fixed mind-set). If studies using passive control groups have the largest effects, this suggests that exposure to treatments might drive the effect rather than growth-mind-set interventions per se. Alternatively, if growth mind-sets are beneficial for academic achievement and fixed mind-sets are detrimental, we should see the largest effect when the comparison group is a fixed-mind-set condition. We examined the type of intervention to test whether interactive (e.g., saying-is-believing) interventions are more effective than passive interventions. The number of intervention sessions was examined as a continuous variable to test whether there is a linear additive effect of intervention exposure. We included mode of intervention (computerized, in person, reading materials, combination) to test whether certain modalities are more effective than others. For interventions at least partially administered in person, we further classified whether administers were teachers, researchers, or both. We include intervention context (integrated in the classroom, outside regular classroom activities) because some researchers have suggested that mind-set interventions might be context dependent (Yeager & Walton, 2011).

We examined whether studies included a manipulation check and whether the manipulation check was successful. We included these measures because if mind-set interventions are a scalable treatment, we should expect most manipulation checks to be successful, and if mind-set interventions are generally effective, we would expect null results only when manipulation checks are unsuccessful.

We also investigated factors related to the measure of academic achievement: intervention-achievement measure interval and type of academic achievement measure. If mind-set interventions are susceptible to the fadeout effect (Protzko, 2015), we should expect stronger effects the shorter the intervention-achievement measure interval. In contrast, if mind-set interventions interact with recursive processes (Yeager & Walton, 2011) the effects should be sustained (or enhanced) with additional time. Finally, we tested whether publication bias is problematic within the mind-set intervention literature.

Method

As with Meta-Analysis 1, we designed the meta-analysis and report the results in accordance with the PRISMA statement (Moher et al., 2009).

Inclusion criteria, literature search, and coding. The criteria for including a study in Meta-Analysis 2 were as follows:

- A growth mind-set treatment, henceforth *intervention*, where the primary goal was to increase students' belief that one or more human attributes (e.g., intelligence) can improve with effort was administered directly to students.
- A control group (active, passive, or fixed-mind-set condition) was included.
- A measure of academic achievement—course exam (e.g., midterm exam), course grade, average of course grades (e.g., GPA), or standardized test performance—was collected.
- An effect size reflecting the difference between the mind-set intervention group and the control group on one or more measures of academic achievement after the intervention was reported, or enough information was provided to compute this effect size.
- The methods and results were in English.

Some studies administered combined interventions, for example, a growth-mind-set intervention immediately followed by another intervention to a single group before measuring academic achievement. We do not include these effects because we cannot know the extent to which the mind-set content is contributing to the effect. Some studies reported results including students who were randomly assigned to the mind-set intervention condition, but did not receive the mind-set intervention. We included effect sizes reflecting only the difference between students who received the mind-set intervention and controls. We excluded two effect sizes, $d = -0.87$ and $d = -0.65$, from two studies conducted by Mendoza-Denton, Kahn, and Chan (2008). Mendoza-Denton and colleagues (2008) designed these studies to reverse any positive effects from a growth-mind-set intervention (relative to a fixed-mind-set condition) when participants received a stereotype lift. They found that students in a fixed-mind-set condition performed better on an academic achievement measure than students in a growth-mind-set condition when experiencing stereotype lift. Presumably, this occurred because fixed mind-sets reinforce the fixedness of stereotyped group differences, which ameliorated self-doubt, anxiety, and other disruptive processes (Mendoza-Denton et al., 2008; see also Walton & Cohen, 2003). These effect sizes were excluded because their inclusion in the meta-analysis could suppress an overall positive effect.

We identified 29 studies that met all the inclusion criteria. We coded each study and the measures collected in it for reference information, student characteristics, methodological characteristics, and results (the data file is available at osf.io/453ds). These studies included 38 independent samples, with 43 effect sizes and a total sample size of 57,155 students. As with Meta-Analysis 1, we adjusted for dependent samples using a method based on that of Cheung and Chan (2004, 2008). For 8 studies, assignment to condition occurred at the classroom or school level rather than at the student level. In these cases, the assumption of independence is violated, and calculations of variance are inappropriately small if not adjusted. Artificially small variances increase the chance of a Type 1 error, and in the case of meta-analyses, extend too much weight to those effect sizes. In these cases, we adjusted (increased) the variance associated with their effect sizes by taking into account the design effect (i.e., multiplying Kish's, 1965, *deff* formula, using the typical intraclass correlation for school effects, $\rho = .10$; Hox, 1998, with the student-level variance) to find the operating variance. For a list of studies included in the meta-analysis, see the Supplemental Material available online. For additional characteristics of Meta-Analysis 2, see Table 2.

Effect sizes. To measure the magnitude of the effectiveness of the intervention, we used Cohen's d as the measure of effect size. Ideally, we would have estimated the difference in gain scores between the treatment and control groups (e.g., Melby-Lervåg, Redick, & Hulme, 2016). However, only a third of the studies provided enough information to calculate this difference. Therefore, except when a study reported a significant pretest difference, we use the standardized mean difference posttreatment scores, which could cause a bias in the effect sizes (see e.g., Melby-Lervåg et al., 2016). Positive Cohen's ds indicated that the group receiving a growth-mind-set intervention performed higher on a measure of academic achievement than students in the control group.

Potential moderators

Student factors. As with Meta-Analysis 1, there were three levels of developmental stage: children (primary school students), adolescents (middle school, junior high school, and high school students), and adults (e.g., postsecondary students). Studies that included students in multiple categories (e.g., elementary school students junior high school students in a single sample) were not included in this moderator analysis.

As with Meta-Analysis 1, there were three levels of academic risk status: high (at risk of failing; e.g., students who previously failed courses), moderate (facing a situational challenge; e.g., transitioning to a new school, a member of a stereotyped group under a stereotype threat manipulation), and low (no indicators that students were at risk). Each sample was categorized on the basis of the majority (> 50%) of the students in the sample. If we could obtain separate effect sizes for each subsample in a study based on risk level, we did so and entered those effects as independent samples. For two studies contributing three effect sizes, we could not obtain separate

Table 2. Descriptive Characteristics for Meta-Analysis 2

Study characteristic	Number of effect sizes (k = 43)	Number of participants (N = 57,155)
Developmental stage[a]		
Children	2	181
Adolescents	27	48,991
Adults	13	7,871
Academic risk status[a]		
Low	17	3,801
Moderate	18	8,664
High	5	1,960
Socioeconomic status[a]		
Low	7	577
Not low	8	4,596
Control group		
Active	26	11,365
Passive	11	45,267
Fixed-mind-set condition	6	523
Intervention type		
Passive	13	1,355
Feedback	1	1,589
Interactive	29	54,211
Mode of intervention		
Computerized training	16	11,581
Reading material	8	1,441
In-person training[b]	14	43,681
By teachers	7	43,141
By researchers	5	649
By both teachers and researchers	5	162
Other (e.g., business partners)	2	181
Both computerized and in-person training	5	452
Intervention context		
Integrated into regular classroom activities	5	2,057
Outside regular classroom activities	38	55,098
Intervention-achievement measure interval[c]		
Immediate (same session)	5	533
Short interval	32	56,180
Long interval	4	292
Manipulation check[d]		
Not included	15	44,484
Included	28	12,671
Significant increase after intervention		
No	13	7,409
Yes	15	5,262
Mind-set type		
Intelligence	33	54,002
Other attribute (e.g., math ability)	10	3,153
Academic achievement measure		
Course exam	3	628
Course grade	4	2,083
Average of course grades (i.e., grade point average)	15	10,564
Standardized test	21	43,880
Laboratory measures	12	889
Publication status		
Published	25	6,180
Unpublished	18	50,975

(continued)

[a]For this characteristic, one or more effect sizes were excluded because this information was not provided in the study.
[b]For this characteristic, the subgroup effect sizes and sample sizes sum to the total sample size for in-person training plus both computerized and in-person training.
[c]For this characteristic, two effect sizes were excluded because the study collected data on grades after a short and long interval but did not identify which of these was reported in the results.
[d]Measures of pre- and postintervention mind-set were included in the study.

effect sizes for each subsample but could obtain a separate effect size for an at-risk subsample (i.e., we obtained the full sample effect size and an at-risk subsample's effect size). In these cases, when examining academic risk status, we report the results in two models: one with the full-sample effect sizes and one where the full-sample effect sizes are replaced with the at-risk students' effect sizes. As with Meta-Analysis 1, we did not code ethnic minorities or women as at risk unless they were under a stereotype threat manipulation.

As with Meta-Analysis 1, there were two levels of SES: low SES (e.g., students qualified for reduced-price lunch) and not low SES (i.e., middle-class or higher). Each study was categorized on the basis of the majority (> 50%) of the students in the sample. Studies not reporting student-level SES were not included in the SES moderator analysis.

Control and intervention method factors. There were three levels of control group type: active control (i.e., placebo control), passive control (e.g., no contact control), and fixed-mind-set condition (i.e., students in the comparison group were given a fixed-mind-set intervention). Students in active (placebo) control groups engaged in similar activities and amounts of contact with administrators but without the content of a hypothesized effective treatment. Active controls did not consist of other treatments designed to be effective in improving academic achievement. When multiple control groups were used in a study, we used the active control whenever possible because this meets a higher scientific standard. If an active control was not available, we used the passive control over a fixed-mind-set condition comparison because a fixed-mind-set condition is often theorized to lower academic achievement (thus, in these cases, it is unclear whether the growth mind-set improves academic achievement or whether the fixed mind-set lowers it).

Intervention type has three levels: passive (students read a document or watch a video on how human attributes are malleable), feedback (students are given feedback on their performance in terms of growth mind-set), and interactive (e.g., participants read materials and then write an essay about how intelligence can be developed or participate in an in-class discussion). If passive and feedback interventions are as effective as interactive interventions, this suggests that effective interventions can be implemented with few resources and with a light touch (see Yeager, Walton, & Cohen, 2013, for a discussion of stealthily implementing interventions).

Intervention length was a continuous variable based on the number of intervention sessions. If intervention effectiveness increases with the number of intervention sessions, then this suggests a positive dose-response relationship. In contrast, if intervention effectiveness decreases with the number of intervention sessions, this could be due to students perceiving the repetition as a message that they need help, undermining the credibility of the growth-mind-set intervention (see Yeager et al., 2013).

Mode of intervention had four levels: computerized training (e.g., Brainology computer program), reading mind-set materials only (e.g., reading how intelligence can change with effort), in-person training (structured discussion or lecture), or a combination of modes (i.e., computerized training and in-person training). Additionally, for in-person training, we examined who administered the intervention. There were three levels to this moderator: researcher, teacher, or both.

Intervention context had two levels: integrated into regular classroom activities (e.g., teacher provides mind-set feedback or fosters discussion of mind-set in class) or outside regular classroom activities. We also examined whether studies included a manipulation check of the mind-set intervention. To be coded as having a manipulation check, the study needed to include a pre- and postintervention measure of mind-set. Of the studies that included a manipulation check, we examined how many were successful.

Academic-achievement-measure-related factors. There were three levels of intervention-achievement measure interval: immediate (within the same session as a mind-set intervention), short interval (within 4 months [approximately a semester's time] of the mind-set intervention), and long interval (longer than 4 months after the mind-set intervention was administered). When studies included measures of academic achievement at multiple time points following the intervention, we used the longest interval available within the same academic term as the intervention. For example, if a study examined students' academic achievement on a course exam a week after the intervention and again on the course final exam 2 months after the intervention, we used the effect size for the final exam. The longest interval within the same academic term was chosen when available to give the intervention time to positively affect study habits and response to failure within the same academic context as the intervention. However, it is also possible that longer intervals, regardless of whether the academic term changes, will increase the effect. Thus, when conducting the intervention-academic achievement interval moderator analysis, we also analyzed the effect sizes from the longest interval between the intervention and the measure of academic achievement regardless of whether the measure of academic achievement occurred in a subsequent term.

As with Meta-Analysis 1, the measure of academic achievement had four levels: course exam (e.g., midterm exam), course grade (e.g., math grade), grade average (e.g., GPA), and standardized test (e.g., Iowa Test of Basic Skills). We also examined this moderator excluding laboratory measures of academic achievement (e.g., administering practice problems from the GRE) that are not part of a student's academic record.

Meta-analytic procedure. The meta-analysis involved four steps. The first step was to obtain the standardized mean difference (Cohen's d) in academic achievement between students who received a growth-mind-set intervention and students in the control group, along with their sampling error variances. The second step was to search for extreme values. There were two outliers—effect sizes whose residuals had z scores of 3 or greater (ds = −0.9554 and 1.5053); we Winsorized these values to z scores equaling 2.99 (ds = −0.9050 and 1.0960, respectively). The third step was to use random-effects meta-analysis modeling, which assumes meaningful differences across studies, to estimate the meta-analytic mean distribution of effects and heterogeneity in the effect sizes, and then to test whether some of the heterogeneity was predictable from moderator variables using mixed-effects meta-analysis modeling. The final step was to perform publication bias analyses.

As with Meta-Analysis 1, we used the Comprehensive Meta-Analysis (Version 2; Borenstein et al., 2005) software package to conduct the meta-analyses. We used Comprehensive Meta-Analysis, the R Project for Statistical Computing (www.r-project.org), the p-curve web application (http://www.p-curve.com/app), and the *p*-uniform web application (https://rvanaert.shinyapps.io/p-uniform) to conduct the publication bias analyses. (See also the Supplemental Material.)

Results

The model consists of 43 effect sizes. Effect sizes are weighted by the inverse of the variance, which includes both between-studies variance and within-study variances. Within-study variance accounts for the sample sizes, such that smaller studies are given less weight, while larger studies are given more weight in the model. The majority of effect-size–associated adjusted *N*s are ≥ 90. The mean adjusted *N* associated with this model's effect sizes is 1,664.

Figure 4 shows that 37 of the 43 effect sizes (86%) are not significantly different from zero. One effect size is significantly different from zero but negative, indicating that students receiving a growth-mind-set intervention had significantly worse academic achievement than students in the control conditions. The remaining 5 effect sizes (12%) are significantly different from zero and positive, indicating that students receiving a growth-mind-set intervention had significantly greater academic achievement than students in the control groups.

The meta-analytic average standardized mean difference (i.e., the average of various population effects) in academic achievement between students receiving a growth-mind-set intervention and students in control groups is $\bar{d}$ = 0.08, 95% CI = [0.02, 0.14], p = .010. When the original outlying effect sizes were entered in lieu of the Winsorized effect sizes, the overall effect did not differ ($\bar{d}$ = 0.08, 95% CI = [0.02, 0.15], p = .010), nor did the pattern of results for the moderator analyses, with the exception that the effect for unpublished studies was no longer significant (see the Supplemental Material for complete results with outliers).

As illustrated by the I^2 statistic, which specifies the percentage of the between-studies variability in effect sizes that is due to heterogeneity rather than random error, there was a medium amount of heterogeneity in the effect sizes, I^2 = 43.15 (τ^2 = .010), indicating that the true effect of a given study could be somewhat lower or higher than the meta-analytic average. We investigated the source of this heterogeneity through the moderator analyses reported next.

Moderator analyses. Williams (2012) recommended at least five cases per subgroup to perform a moderator analysis in meta-analyses. We therefore did not include any groups with fewer than five effect sizes when conducting moderator analyses. See Table 2 for the number of effect sizes per group.

Student factors. The developmental stage of the students was not a significant moderator, $Q(1) < 0.01$, p = .999. Only two effect sizes associated with children were available. For one effect size, sample age information was unavailable. These three effect sizes were not included in this analysis. Growth-mind-set intervention did not significantly improve academic achievement relative to controls either for adolescents, $\bar{d}$ = 0.08, 95% CI = [−0.01, 0.17], p = .090, or for adults, $\bar{d}$ = 0.08, 95% CI = [−0.02, 0.17], p = .123.

Academic at-risk status was not a significant moderator, $Q(2)$ = 0.67, p = .715. Sample descriptions were unavailable for three effect sizes, and thus these effect sizes were not included in this analysis. Growth-mind-set intervention did not significantly improve academic achievement relative to controls for low-risk students, $\bar{d}$ = 0.06, 95% CI = [−0.01, 0.12], p = .109; for moderately at-risk students, $\bar{d}$ = 0.08, 95% CI = [−0.04, 0.19], p = .177; or for highly at-risk students, $\bar{d}$ = 0.17, 95% CI = [−0.11, 0.45], p = .231. In two studies (with 3 effect sizes), authors provided separate effect sizes for subsamples of students who were high risk. When replacing these three effect sizes and variances of the full samples (where the majority were not high risk) with these subsamples, the moderator remains nonsignificant, $Q(2)$ = 2.18, p = .335. However, the high-risk group (8 effect sizes) then demonstrated a borderline significant effect of growth-mind-set intervention, $\bar{d}$ = 0.19, 95% CI = [0.02, 0.36], p = .031. Low-risk students and moderate-risk students did not benefit from a growth-mind-set intervention, $\bar{d}$ = 0.05, 95% CI = [−0.02, 0.12], p = .162, and $\bar{d}$ = 0.09, 95% CI = [−0.04, 0.23], p = .162, respectively.

SES was a significant moderator, $Q(1)$ = 4.76, p = .029. Student-level SES was not reported for 28 effect sizes, and thus these effect sizes were not included in this analysis. Growth-mind-set intervention did not improve middle-class and upper-class students' academic achievement, $\bar{d}$ = 0.03, 95% CI = [−0.06, 0.11], p = .538. However, for those from low-SES households (7 effect sizes), academic achievement was significantly higher for students who received growth-mind-set interventions relative to controls, $\bar{d}$ = 0.34, 95% CI = [0.07, 0.62], p = .013.

Control and intervention-related factors. Control-group type was not a significant moderator, $Q(2)$ = 2.96, p = .228. Academic achievement was similar between students who received a growth-mind-set intervention and students who received a fixed-mind-set condition, $\bar{d}$ = 0.27, 95% CI = [−0.05, 0.59], p = .100. There was also no effect when the control group was passive, $\bar{d}$ = 0.02, 95% CI = [−0.05, 0.10], p = .522. A borderline significant difference was observed when the control group was an active control (i.e., placebo control), $\bar{d}$ = 0.08, 95% CI = [0.01, 0.16], p = .034.

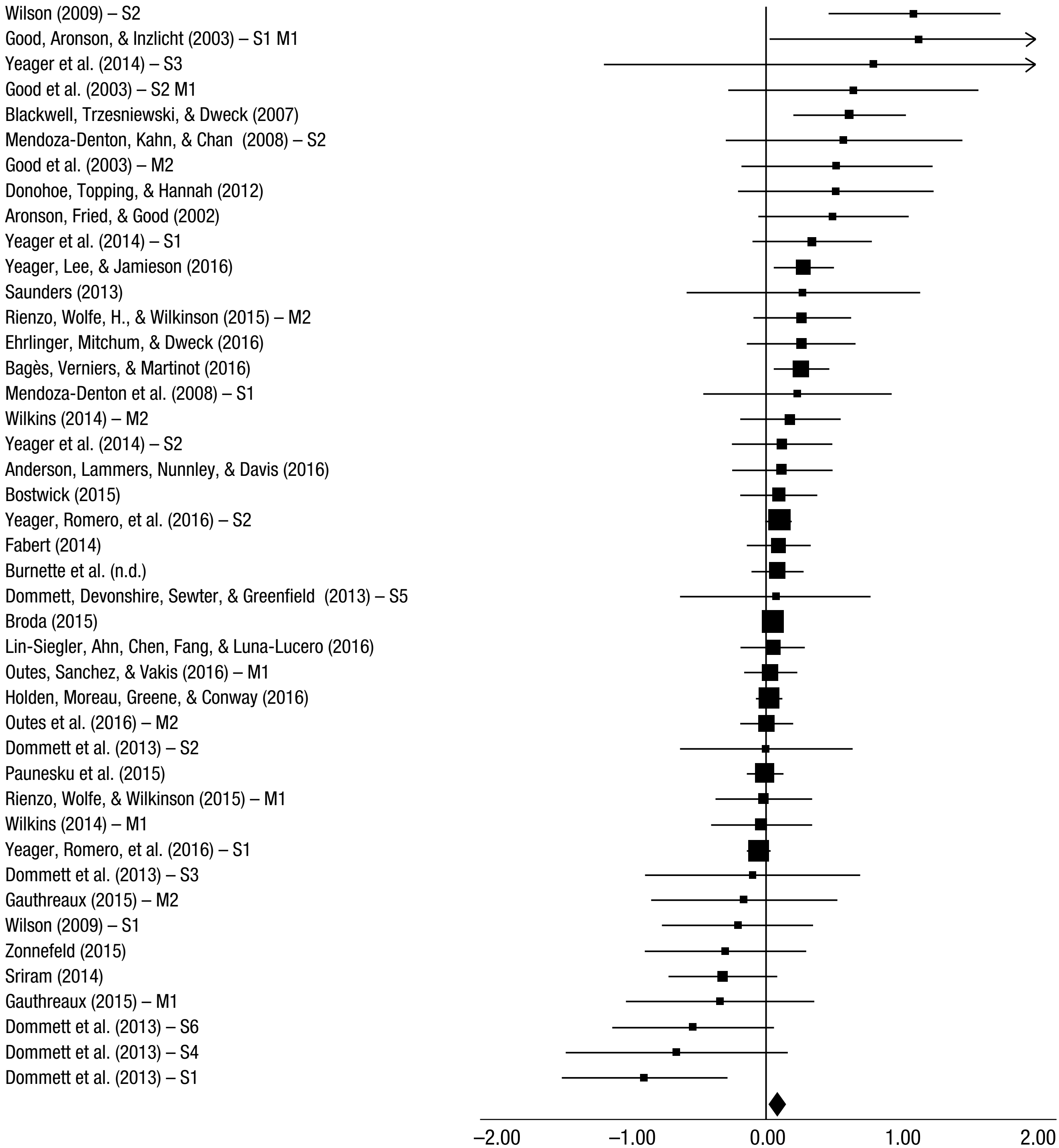

Fig. 4. Standardized mean differences (Cohen's *d*s) in academic achievement between students receiving a growth-mind-set intervention and students in the comparison group. Cohen's *d*s (squares) and 95% confidence intervals (error bars) are displayed for all effects entered into Meta-Analysis 2. The size of the square represents the effect size's meta-analytic weight. The diamond on the bottom row represents the meta-analytically weighted mean Cohen's *d*. For studies with multiple independent samples, the result for each sample (S1, S2, etc.) is reported separately. Similarly, for studies with multiple measures, the result for each measure (M1, M2, etc.) is reported separately. Multiple measures were adjusted for dependency. See the Supplemental Material available online for full details of all references.

Intervention type was not a significant moderator, $Q(1) = 0.58$, $p = .447$. Only one effect size used feedback (weekly growth mind-set feedback with students' quiz grades) as the manipulation, and thus this effect size was removed from this moderator analysis. The effectiveness of a growth-mind-set intervention on academic achievement was not significant when the intervention was passive (e.g., reading about growth mind-set without writing a reflection), $\bar{d} = 0.02$, 95% CI = [−0.16, 0.19], $p = .852$, but demonstrated effectiveness when the intervention was interactive (e.g., reading about growth mind-set and then writing a reflection), $\bar{d} = 0.09$, 95% CI = [0.02, 0.16], $p = .011$.

Intervention length was not a significant moderator, $Q(1) = 0.12$, $b = -0.005$, 95% CI = [−0.03, 0.02], $p = .734$. The number of sessions ranged from 1 to 10. Increasing the number of growth-mind-set-intervention sessions neither increased nor decreased the impact on academic achievement.

Mode of intervention was a significant moderator, $Q(3) = 9.33$, $p = .025$. Growth-mind-set interventions were not effective when administered via computer programs, $\bar{d} = 0.03$, 95% CI = [−0.03, 0.08], $p = .409$; in person, $\bar{d} = 0.06$, 95% CI = [−0.12, 0.25], $p = .517$; or via a combination of modes, $\bar{d} = 0.27$, 95% CI = [−0.04, 0.59], $p = .092$. The intervention was effective when students read growth-mind-set materials, $\bar{d} = 0.20$, 95% CI = [0.09, 0.30], $p < .001$. Post hoc follow-up tests were conducted to examine the source of this heterogeneity with an adjusted α of .008 (.05/6). The follow-up tests revealed that mind-set interventions administered via reading materials were significantly more effective than when administered via computer programs, $Q(1) = 7.75$, $p = .005$. No other pairwise comparisons were significant, all *p*s > .133.

When interventions were administered in person (solely or as part of a combination), growth-mind-set interventions remained ineffective regardless of whether the intervention was administered by a teacher, $\bar{d} = -0.01$, 95% CI = [−0.14, 0.12], $p = .882$; a researcher, $\bar{d} = 0.34$, 95% CI = [−0.14, 0.82], $p = .167$; or both, $\bar{d} = 0.27$, 95% CI = [−0.23, 0.77], $p = .296$; $Q(2) = 2.80$, $p = .246$. Two effect sizes were associated with administrators other than a teacher or researcher and were thus removed from this analysis.

The context in which the intervention was implemented was not a significant moderator, $Q(1) = 2.52$, $p = .112$. Growth-mind-set interventions were not effective when the intervention was integrated into regular classroom activities, $\bar{d} = -0.14$, 95% CI = [−0.43, 0.14], $p = .327$. However, when the interventions were administered outside regular classroom activities, the effect was significant, $\bar{d} = 0.09$, 95% CI = [0.03, 0.16], $p = .003$.

Fifteen of the 43 effect sizes (35%) were associated with studies that did not report pre- and postintervention measures of mind-set to test whether the growth-mind-set intervention effectively increased growth mind-set (i.e., no manipulation checks). Interestingly, the effect of a growth-mind-set intervention was significant when no manipulation check was administered, $\bar{d} = 0.18$, 95% CI = [0.05, 0.31], $p = .005$, but not significant for studies that employed a manipulation check, $\bar{d} = 0.04$, 95% CI = [−0.03, 0.10], $p = .249$. The difference between these two groups of studies was borderline significant, $Q(1) = 3.95$, $p = .047$.

For the 28 effect sizes associated with studies that did employ a manipulation check, almost half (46%) failed to observe a significant difference between pre- and postintervention measures of mind-set. The effectiveness of the growth-mind-set intervention was borderline significant only when the manipulation check failed, $\bar{d} = 0.05$, 95% CI = [0.001, 0.09], $p = .044$, indicating that students' growth mind-sets had not changed following the intervention. Growth-mind-set interventions were not effective for the studies where the manipulation check succeeded, $\bar{d} = 0.02$, 95% CI = [−0.11, 0.15], $p = .771$; $Q(1) = 0.18$, $p = .672$.

Factors related to academic achievement measures. When using effect sizes associated with the greatest amount of time between the intervention and measure of academic achievement within the same semester, if available, the interval between the growth-mind-set intervention and the measure of academic achievement was not a significant moderator, $Q(1) = 2.41$, $p = .121$. Two effect-size intervals were ambiguous. Only four effect sizes measured solely long-term (> 4 months) academic achievement following the interventions. These 6 effect sizes were excluded from this analysis. The effectiveness of growth-mind-set interventions was not significant regardless of whether academic achievement was measured within the same session, $\bar{d} = 0.35$, 95% CI = [−0.03, 0.72], $p = .070$, or within 4 months of the intervention, $\bar{d} = 0.05$, 95% CI = [−0.01, 0.11], $p = .126$.

Nine effect sizes were associated with studies that measured academic achievement at two time points, once within the same semester as the intervention and once following that semester. For seven of these effects, the second measure was administered more than 4 months after the intervention. For two of these effects, the second measure was administered following the semester of the intervention, but within 4 months of the intervention. When replacing the seven short-term-interval effect sizes with their long-term-interval counterparts, replacing the two short-term effect sizes with their longest interval (though still short-term) counterparts, and adding in the four long-term-interval effect sizes that were previously excluded because there were fewer than five, we found that interval was still not a significant moderator, $Q(2) = 3.92$, $p = .141$. The two effect sizes mentioned previously where the interval was ambiguous were not included in this analysis. Short-term academic achievement (within 4 months of the intervention) remained nonsignificant, $\bar{d} = 0.05$, 95% CI = [−0.002, 0.10], $p = .057$. Long-term academic achievement (measured more than 4 months after the intervention) was also nonsignificant, $\bar{d} = 0.19$, 95% CI = [−0.02, 0.39], $p = .072$.

The type of academic achievement measure was not a significant moderator, $Q(1) = 0.03$, $p = .862$. Only three effect sizes were associated with a course exam grade, and only four effect sizes were associated with a course grade. We did not include these seven effect sizes in this moderator analysis. Growth-mind-set interventions were borderline significant when the measure of academic achievement was GPA, $\bar{d} = 0.07$, 95% CI = [0.002, 0.14], $p = .045$. Growth-mind-set interventions were not significant when the measure of academic achievement was performance on a standardized test, $\bar{d} = 0.09$, 95% CI = [−0.07, 0.24], $p = .276$.

Twelve effect sizes were from standardized tests (or portions of standardized tests) administered by researchers in a laboratory setting (coded as standardized test). Excluding these 12 effect sizes did not change the overall results, $\bar{d} = 0.06$, 95% CI = [0.01, 0.11], $p = .012$, or the pattern of results for the academic achievement moderator. Measure of academic achievement remained a nonsignificant moderator, $Q(1) = 0.07$, $p = .796$, and growth-mind-set interventions remained ineffective when the measure of academic achievement was performance on an actual standardized test, $\bar{d} = 0.09$, 95% CI = [−0.05, 0.24], $p = .213$.

Publication bias analyses. We conducted the same three types of publication bias analyses as in Meta-Analysis 1.

Moderator analysis. The median sample size associated with unpublished studies was 270 (compared with 66 for published studies). A moderator analysis revealed that the 18 effect sizes associated with unpublished studies ($\bar{d} = 0.05$, 95% CI = [0.004, 0.10], $p = .032$) were not significantly different from the 25 effect sizes associated with published studies ($\bar{d} = 0.11$, 95% CI = [0.002, 0.22], $p = .045$), $Q(1) = 1.02$, $p = .313$.

p*-curve analysis.* Only four statistically significant results were available to be included in the primary analysis (*p*-curve excludes unpublished results and nonsignificant results). We performed a simulation in R (www.r-project.org; Simmons & Simonsohn, 2017) to assess the power of *p*-curve to detect right-skew for four studies, each with 16.5% power. The estimated power of the *p*-curve analysis to detect right-skew on the basis of this simulation was 18.2%. Therefore, the results of the *p*-curve analyses are inconclusive. See the Supplemental Material for the results, additional figures, and an index of the effect sizes that were entered into the analyses.

Egger's regression. During our search, we became aware of multiple unpublished studies for which we could not access the methods or results. As one example, when we requested results for an unpublished intervention discussed in a presentation, the researcher declined to provide this information on the grounds that replication attempts had failed. The researcher also declined to provide access to the results of the failed replications. Despite being aware of missing studies, we found that the funnel plot was approximately symmetrical, suggesting that our meta-analysis was unaffected by missing studies with weaker-than-average effect sizes, $B_0 = 0.38$, 95% CI = [−0.20, 0.97], $t(41) = 1.33$, $p = .192$.

Discussion

Some researchers have claimed that mind-set interventions can "lead to large gains in student achievement" and have "striking effects on educational achievement" (Yeager & Walton, 2011, pp. 267 and 268, respectively). Overall, our results do not support these claims. Mind-set interventions on academic achievement were nonsignificant for adolescents, typical students, and students facing situational challenges (transitioning to a new school, experiencing stereotype threat). However, our results support claims that academically high-risk students and economically disadvantaged students may benefit from growth-mind-set interventions (see Paunesku et al., 2015; Raizada & Kishiyama, 2010), although these results should be interpreted with caution because (a) few effect sizes contributed to these results, (b) high-risk students did not differ significantly from non-high-risk students, and (c) relatively small sample sizes contributed to the low-SES group.

The results do not support the claim that mind-set interventions benefit both high- and low-achieving students (e.g., see Mindset Scholars Network[4]). Mind-set interventions are relatively low cost and take little time, so there may be a net benefit for students' academic achievement. However, there may be a detriment relative to fixed-mind-set conditions when students are confident in their abilities (Mendoza-Denton et al., 2008). Regardless, those seeking more than modest effects or effects for all students are unlikely to find them. To this end, policies and resources targeting all students might not be prudent.

Regarding methodological moderators, interactive interventions produced a significant effect in line with mind-set theory (Yeager et al., 2013). However, other results were confusing. For example, there was no significant difference between students in growth-mind-set versus fixed-mind-set conditions or when the treatment group was passive—the effect was significant only when compared with active controls. As another example, the effect was significant for studies that did not report manipulation checks while nonsignificant for studies with manipulation checks. Further, of studies that reported manipulation checks, almost half failed, suggesting that the interventions had no impact on students' mind-sets. Most surprising, the effect was significant when the manipulation checks failed but null when the manipulation checks succeeded. This suggests that "successful" interventions may not be attributable to students' mind-sets. Manipulation checks are critical for establishing causal inferences (Alferes, 2012).

General Discussion

Mind-sets and their implications for academic achievement have received substantial attention from the media (e.g., PBS, Time, NPR; see Paul, 2013; Eisenberg, 2005; Smith, 2014, respectively), funding agencies, educators, and government institutions. For example, in 2013, the White House convened a special meeting entitled "Excellence in Education: The Importance of Academic Mindsets." Boaler (2013) summarized the impact as the "mindset revolution that is reshaping education."

Part of the reshaping effort has been to make funding mind-set research a "national education priority" (Rattan et al., 2015, p. 723) because mind-sets have "profound effects" on school achievement (Dweck, 2008, para. 2). Our meta-analyses do not support this claim. Effect sizes were inconsistent across studies, but most analyses yielded small (or null) effects. Overall, the first meta-analysis demonstrated only a very weak relationship between mind-sets and academic achievement. Similarly, the second meta-analysis demonstrated only a very small overall effect of mind-set interventions on academic achievement.

However, not all mind-set research makes broad claims. Some research focuses on specific tenets of the theory regarding how mind-sets affect individuals facing challenges, hypothesizing effects only for specific groups of students. Some subgroup results from the present meta-analyses supported these hypotheses, such as the significant effects for academically high-risk students and low-SES students. Other subgroup results did not support these hypotheses, such as null results for students facing situational challenges and adolescents. Still other subgroup results suggest that standards are needed for implementing intervention studies and interpreting the results.

Moving forward, researchers interested in mind-sets' effects on academic achievement should institute manipulation checks to ensure that mind-set interventions are influencing students' mind-sets. If mind-set manipulations are not demonstrating an influence on students' mind-sets (as was found in nearly half the studies including manipulation checks), then the mechanism affecting any observed change in achievement is either due to chance or due to mediating variables. Additionally, while the results that supported mind-set theory were not strong, it is possible that unmeasured factors are suppressing effects or that imperfect control of the intervention in the classroom buffers the effects (Yeager & Walton, 2011). Alternatively, mind-set interventions might need to be combined with other interventions to increase effectiveness. From a theoretical perspective, further investigations into potential mediators and moderators might yield important discoveries about the nature of human beliefs, the role of educational interventions, or both.

However, from a practical perspective, resources might be better allocated elsewhere than mind-set interventions. Across a range of treatment types, Hattie, Biggs, and Purdie (1996) found that the meta-analytic average effect size for a typical educational intervention on academic performance is 0.57. All meta-analytic effects of mind-set interventions on academic performance were < 0.35, and most were null. The evidence suggests that the "mindset revolution" might not be the best avenue to reshape our education system.

Action Editor

D. Stephen Lindsay served as action editor for this article.

Author Contributions

V. F. Sisk developed the study concept with input from J. L. Butler and B. N. Macnamara. V. F. Sisk and B. N. Macnamara developed the methodologies with input from J. L. Butler. V. F. Sisk, A. P. Burgoyne, and J. Sun performed effect-size data collection with input from B. N. Macnamara and J. L. Butler. B. N. Macnamara and V. F. Sisk coded moderators and conducted calculations for Meta-Analysis 1. V. F. Sisk and A. P. Burgoyne coded moderators and conducted calculations for Meta-Analysis 2. B. N. Macnamara performed the data analyses. A. P. Burgoyne and B. N. Macnamara performed the publication bias analyses. B. N. Macnamara drafted the introduction, Method, Results, and Discussion sections of the manuscript. B. N. Macnamara and A. P. Burgoyne drafted the results for the publication bias analysis. V. F. Sisk, A. P. Burgoyne, and J. L. Butler provided critical revisions. All authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617739704

Open Practices

All data have been made publicly available via the Open Science Framework and can be accessed at osf.io/453ds. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617739704. This article has received the badge for Open Data. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

Notes

1. See, for example, the PERTS press kit (www.perts.net/press_kit).
2. See the U.S. Department of Education's Press Releases page (www.ed.gov/news/press-releases/us-department-education-announces-first-ever-skills-success-grants-and-initiative-support-learning-mindsets-and-skills).
3. See the website of the Institute of Education Sciences (ies.ed.gov/funding/grantsearch/details.asp?ID=1728).
4. A relevant quote from the "Improving Student Outcomes and Expanding Educational Opportunity" section on the Mindset Scholars Network's FAQ page regarding this claim is, "Learning mindsets have been shown to be beneficial at every level: from students struggling academically in middle school to undergraduates at highly selective universities" (http://mindsetscholarsnetwork.org/learning-mindsets/faq/).

References

Alferes, V. R. (2012). *Methods of randomization in experimental design*. Thousand Oaks, CA: SAGE.

Becker, G. (1986). Correcting the point-biserial correlation for attenuation owing to unequal sample size. *The Journal of Experimental Education*, *55*, 5–8.

Bishop, D. V. M., & Thompson, P. A. (2016). Problems in using *p*-curve analysis and text-mining to detect rate of *p*-hacking and evidential value. *PeerJ*, *4*, Article e1715. doi:10.7717/peerj.1715

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*, 246–263. doi:10.1111/j.1467-8624.2007.00995.x

Boaler, J. (2013). Ability and mathematics: The mindset revolution that is reshaping education. *Forum*, *55*, 143–152.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive Meta-Analysis (version 2) [Computer software]. Englewood, NJ: Biostat.

Burnette, J. L., & Finkel, E. J. (2012). Buffering against weight gain following dieting setbacks: An implicit theory intervention. *Journal of Experimental Social Psychology*, *48*, 721–725. doi:10.1016/j.jesp.2011.12.020

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, *139*, 655–701.

Cheung, S. F., & Chan, D. K.-S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology*, *89*, 780–791.

Cheung, S. F., & Chan, D. K.-S. (2008). Dependent correlations in meta-analysis: The case of heterogeneous dependence. *Educational and Psychological Measurement*, *68*, 760–777. doi:10.1177/0013164408315263

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences, USA*, *113*, 8664–8668. doi:10.1073/pnas.1608207113

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. East Sussex, England: Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.

Dweck, C. S. (2008, Winter). Brainology: Transforming students' motivation to learn. *Independent School*. Retrieved from https://www.nais.org/magazine/independent-school/winter-2008/brainology/

Dweck, C. S. (2012). Mindsets and human nature: Promoting change in the Middle East, the schoolyard, the racial divide, and willpower. *American Psychologist*, *67*, 614–622. doi:10.1037/a0029783

Dweck, C. S. (2016). The remarkable reach of growth mindsets. *Scientific American Mind*, *27*, 36–41. doi:10.1038/scientificamericanmind0116-36

Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*, 267–285.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629–634.

Eisenberg, D. (2005, November 6). How to help them succeed. *Time*. Retrieved from http://content.time.com/time/magazine/article/0,9171,1126743,00.html

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66*, 99–136.

Hox, J. (1998). Multilevel modeling: When and why. In I. Balderjahn, R. Mathar, & M. Schader (Eds.), *Classification, data analysis, and data highways* (pp. 147–154). Berlin, Germany: Springer.

Hunter, J. E., & Schmidt, F. L. (1990). Dichotomization of continuous variables: The implications for meta-analysis. *Journal of Applied Psychology*, *75*, 334–339. doi:10.1037/0021-9010.75.3.334

Kish, L. (1965). *Survey sampling*. New York, NY: Wiley.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*, 512–534. doi:10.1177/1745691616635612

Mendoza-Denton, R., Kahn, K., & Chan, W. (2008). Can fixed views of ability boost performance in the context of favorable stereotypes? *Journal of Experimental Social Psychology*, *44*, 1187–1193. doi:10.1016/j.jesp.2008.03.005

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, *6*(7), Article e1000097. doi:10.1371/journal.pmed1000097

Paul, A. M. (2013, June 11). *The science of smart: Eight ways of looking at intelligence*. PBS. Retrieved from http://www.pbs.org/wgbh/nova/blogs/secretlife/blogposts/the-science-of-smart-eight-ways-of-looking-at-intelligence/

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, *26*, 784–793. doi:10.1177/0956797615571017

Protzko, J. (2015). The environment in raising early intelligence: A meta-analysis of the fadeout effect. *Intelligence*, *53*, 202–210.

Raizada, R. D., & Kishiyama, M. M. (2010). Effects of socioeconomic status on brain development, and how cognitive neuroscience may contribute to levelling the playing field. *Frontiers in Human Neuroscience*, *4*, Article 3. doi:10.3389/neuro.09.003.2010

Rattan, A., Savani, K., Chugh, D., & Dweck, C. S. (2015). Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspectives on Psychological Science*, *10*, 721–726. doi:10.1177/1745691615599383

Romero, C., Master, A., Paunesku, D., Dweck, C. S., & Gross, J. J. (2014). Academic and emotional functioning in middle school: The role of implicit theories. *Emotion*, *14*, 227–234. doi:10.1037/a0035490

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.

Simmons, J. P., & Simonsohn, U. (2017). Power posing: *P*-curving the evidence. *Psychological Science*, *28*, 687–693.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014a). *P*-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014b). *p*-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, *9*, 666–681.

Smith, T. (2014, March 17). *Does teaching kids to get 'gritty' help them to get ahead?* NPR. Retrieved from https://www.npr.org/sections/ed/2014/03/17/290089998/does-teaching-kids-to-get-gritty-help-them-get-ahead

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343*, Article d4002. doi:10.1136/bmj.d4002

van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2016). Conducting meta-analyses based on *p* values: Reservations and recommendations for applying *p*-uniform and *p*-curve. *Perspectives on Psychological Science*, 11, 713–729.

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*, 456–467. doi:10.1016/S0022-1031(03)00019-2

Williams, R. (2012, May). *Moderator analyses: Categorical models and meta-regression*. Paper presented at the annual Campbell Collaboration Colloquium, Copenhagen, Denmark.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*, 302–314. doi:10.1080/00461520.2012.722805

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*, 267–301. doi:10.3102/0034654311405999

Yeager, D. S., Walton, G., & Cohen, G. L. (2013). Addressing achievement gaps with psychological interventions. *Phi Delta Kappan*, *94*(5), 62–65.