Psychological Bulletin
2014, Vol. 140, No. 4, 1174–1204

0033-2909/14/$12.00 http://dx.doi.org/10.1037/a0036620

Gender Differences in Scholastic Achievement: A Meta-Analysis

Daniel Voyer and Susan D. Voyer
University of New Brunswick

A female advantage in school marks is a common finding in education research, and it extends to most course subjects (e.g., language, math, science), unlike what is found on achievement tests. However, questions remain concerning the quantification of these gender differences and the identification of relevant moderator variables. The present meta-analysis answered these questions by examining studies that included an evaluation of gender differences in teacher-assigned school marks in elementary, junior/middle, or high school or at the university level (both undergraduate and graduate). The final analysis was based on 502 effect sizes drawn from 369 samples. A multilevel approach to meta-analysis was used to handle the presence of nonindependent effect sizes in the overall sample. This method was complemented with an examination of results in separate subject matters with a mixed-effects meta-analytic model. A small but significant female advantage (mean $d = 0.225$, 95% CI [0.201, 0.249]) was demonstrated for the overall sample of effect sizes. Noteworthy findings were that the female advantage was largest for language courses (mean $d = 0.374$, 95% CI [0.316, 0.432]) and smallest for math courses (mean $d = 0.069$, 95% CI [0.014, 0.124]). Source of marks, nationality, racial composition of samples, and gender composition of samples were significant moderators of effect sizes. Finally, results showed that the magnitude of the female advantage was not affected by year of publication, thereby contradicting claims of a recent "boy crisis" in school achievement. The present meta-analysis demonstrated the presence of a stable female advantage in school marks while also identifying critical moderators. Implications for future educational and psychological research are discussed.

Keywords: school achievement, school grades, gender differences, meta-analysis

Much research has focused on gender differences in various areas of intellectual achievement (Halpern, 2012). In fact, reliance on this research often guides policy decisions such as funding for sex-segregated education (Lindberg, Hyde, Petersen, & Linn, 2010).

In reality, much of what we know about gender differences in intellectual achievement comes from various meta-analyses that have summarized and quantified the findings obtained in relevant research. For example, Hyde, Fennema, and Lamon (1990; see also Else-Quest, Hyde, & Linn, 2010) reported that gender differences in mathematics achievement were typically in favor of males,[1] although recent data suggest that the gap is closing (or even disappearing) in this field (Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Lindberg et al., 2010). A male advantage has also been reported for science achievement tests (Hedges & Nowell, 1995), whereas a female advantage is typically reported in reading comprehension (e.g., Hedges & Nowell, 1995; Lynn & Mikk, 2009; Nowell & Hedges, 1998). These findings have essentially become part of the stereotypical view of men and women (Lindberg et al., 2010; Nosek et al., 2009).

Defining Achievement

To our knowledge, all existing meta-analyses examining mathematics, science, and reading achievement have relied either on tests of cognitive abilities or on national test scores as their measures of focus. For example, Nowell and Hedges (1998) examined results obtained in specific data sets in which participants completed a battery of cognitive measures used in a national test (see Table 2 in their article for a list). Such achievement tests have been shown to predict later performance in the classroom (e.g., Anastasi, 1988, for the Scholastic Aptitude Test; Kuncel, Henzltet, & Ones, 2001, for the Graduate Record Examination). Although gender differences follow essentially stereotypical patterns on achievement tests, for whatever reasons, females generally have the advantage on school marks[2] regardless of the material. This gender difference has been known to exist for many years (e.g.,

Daniel Voyer and Susan D. Voyer, Department of Psychology, University of New Brunswick.

This research was made possible by a research grant awarded by the Natural Sciences and Engineering Research Council of Canada (NSERC) to Daniel Voyer. We are indebted to Lucia Tramonte for her advice on some thorny HLM issues and to Kaitlyn Fallow and Kathryn Malcom for their assistance with literature retrieval. This research also owes much to the efficiency of the University of New Brunswick Document Delivery library personnel.

Correspondence concerning this article should be addressed to Daniel Voyer, Department of Psychology, University of New Brunswick, P.O. Box 4400, Fredericton, NB, Canada, E3B 5A3. E-mail: voyer@unb.ca

---

[1] We use the terms *males* and *females* throughout this article to cover a variety of age groups. We use the terms *girls* and *boys* or *women* and *men* when relevant to the age of the specific samples under discussion.

[2] We use the word *marks* throughout this article to refer to the same thing that is often called *grades.* Essentially, we used *marks* to remove the ambiguity inherent in the word *grades,* which could also reflect grade level (Grade 1, Grade 2, etc.).

Goodenough, 1954; Hosseini, 1975; Kimball, 1989) and has persisted in recent years (e.g., McCornack & McLeod, 1988, for college students; Pomerantz, Altermatt, & Saxon, 2002, for children in Grades 4, 5, and 6).

Many attempts have been made to explain the apparent contradiction between what is observed on gender differences with achievement tests and with actual school performance. For example, Wentzel (1991) suggested that school marks reflect learning in the larger social context of the classroom. School marks also require effort and persistence over long periods of time, whereas performance on standardized tests assesses basic or specialized academic abilities and aptitudes at one point in time without social influences. Kimball (1989) and Kenney-Benson, Pomerantz, Ryan, and Patrick (2006) elaborated on factors that distinguish school marks and achievement tests, including the type of learning required, the familiarity level of each of these test settings, the role of anxiety and confidence in performance, the influence of subjective factors, and gender stereotypes. An in-depth coverage of these factors is beyond the scope of the present article. However, existing data for achievement tests are based on meta-analytic findings, whereas research on teacher-assigned school marks has not been examined as a whole to date. Therefore, a systematic meta-analysis examining gender differences in school marks is clearly required to summarize the literature on this topic in the same way that this has been done for achievement tests.

Accordingly, the purpose of the present research was to fill that gap by reporting the results of a meta-analysis of gender differences in scholastic achievement focusing exclusively on school achievement in the form of teacher-assigned school marks. Two questions were considered as part of that purpose. First, are there overall gender differences in school performance? Second, what factors moderate these gender differences? The first question is straightforward as it requires only a global test of significance on a set of retrieved studies. However, the second question requires the a priori identification of potential factors that might moderate the magnitude of the gender differences.

Potential Moderating Variables

The identification of potential moderating variables relies on what researchers have considered important in past research. From this perspective, factors that have been examined both in research on school marks (the focus of the present article) and in work examining tests of achievement and cognitive abilities are relevant.

Age

Age of the participants sampled in a given study is likely the most obvious variable to consider. However, when considering school achievement, age in years is confounded with school level (preschool, elementary, high school, college). From this perspective, age can be considered either as a continuous variable or as a categorical variable. Lindberg et al. (2010), who opted for the categorical approach, found much variability in the magnitude of gender differences depending on whether achievement scores were obtained with preschool, elementary school, middle school, high school, college, or adult samples. It is possible that this categorical approach produced a significant effect of this moderator in their analysis because it captured the existing nonlinearity that the typical linear metaregression approach based on a continuous variable would not detect. Accordingly, a categorical approach is followed here when assessing age. However, one also has to consider that there are studies sampling, for example, university students, but their high school grade point average (GPA) is reported. This discrepancy means that an examination of age in the present analysis has to be further refined. Specifically, the source of the grades is more informative than the mean age reported in the study or the current school level of the participants. Accordingly, the source of the grades was used as a variable in the present analysis.

In their analysis of achievement tests, Lindberg et al. (2010) reported that the male advantage in mathematics achievement tests increased with age, with a peak in high school and a decline in college and adult samples. In contrast, in our analysis of school marks, we predict a decrease in the magnitude of gender differences with sources that reflect different age groups. Specifically, individual studies suggest that a female advantage is found in elementary school (Pomerantz et al., 2002), middle school (Mickelson & Greene, 2006), and high school (McCornack & McLeod, 1988). At the university level, findings are more variable, with some researchers reporting a female advantage (McCornack & McLeod, 1988), others reporting no gender difference (Sulaiman & Mohezar, 2006), and yet others reporting a male advantage (Beaudin, Horvath, & Wright, 1992). Thus, gender differences at the university level should reflect more dilution of the effects.

Course Material

The actual topic on which the school marks were based is another rather obvious variable requiring consideration. GPA is a global score and can thus be considered as a composite measure that might result in somewhat heterogeneous findings depending on the combination of courses that form the score. However, as it is expected that the magnitude of gender differences in school performance should fluctuate as a function of the course material, we attempted to enter separate grades for each material type. Accordingly, we entered a global score such as the GPA in the meta-analytic data set when it was the only one available. Fortunately, in many studies, the authors presented data for specific subject matters so that effect sizes for gender differences in each subject were considered in the analysis. Research examining school performance suggests the expectation that females should outperform males in all subjects (Pomerantz et al., 2002), including global measures.

National Origin

Potential national differences in the magnitude of gender differences in school achievement as well as in test achievement have been examined extensively (Else-Quest et al., 2010). Thus, it is only fitting that this variable should be considered here. However, in view of the potential variability in cross-national findings and in the composition of the final sample of studies, this variable should be seen as exploratory. Therefore, no predictions on possible outcomes are presented at this time.

Year of Publication

Some support exists for the notion that gender differences have decreased in magnitude for some areas of cognitive achievement in recent years (Feingold, 1988; Voyer, Voyer, & Bryden, 1995). Specifically, Feingold (1988) and Voyer et al. (1995) reported a trend for decreasing gender differences in tests of spatial abilities as a function of year of publication. These results have typically been interpreted as reflecting social changes that promote more equality in how children are raised and educated. To determine whether such trends would be found for gender differences in school marks, year of publication is examined in the present analysis.

The inclusion of year of publication as a moderator also has important implications to address popular views on gender differences in school achievement. In particular, the literature presented earlier suggests that many researchers have been aware of a female advantage in school achievement for decades. However, it is interesting that it just received media attention recently, leading to the notion that this is a new phenomenon. Specifically, a 2006 *Newsweek* article (Tyre, 2006) suggested that boys across the United States are falling behind girls in terms of school achievement, whereas, 30 years ago, it was presumably females who were lagging. Unfortunately, no specific references were provided to support these statements. However, this did not prevent more reporting of this so-called boy crisis in various newspapers, magazines, and other media (Rybak Lang, 2011). In fact, this issue has also received attention in many countries (Cappon & Canadian Council on Learning, 2011). Critics such as Rybak Lang (2011), Vail (2006), and Mead (2006) showed skepticism about the notion of crisis, suggesting that males' performance has not declined but rather females' performance has improved. Most of the data (at least in Cappon & Canadian Council on Learning, 2011, and Mead, 2006) are based on relatively recent achievement test scores or enrollment figures, but the school performance data required for a complete test of the boy-crisis claims are lacking. Fortunately, the examination of year of publication as a potential moderator of gender differences in school achievement fills this gap. Essentially, a positive relation between year and magnitude of the effect would support the notions underlying the claims of a boy crisis by showing that gender differences have either changed in direction or increased in magnitude. In contrast, a nonsignificant or negative relation would allow us to reject the central claims inherent in the boy crisis. Therefore, this important question is examined with the inclusion of year of publication as a moderator in the present analysis.

Racial Composition of the Sample

In her discussion of the boy crisis, Mead (2006) also pointed out that achievement gaps are actually larger when racial origin is examined as opposed to gender in United States samples. For example, the graduation rate is much lower for students who are Hispanic or Black when compared to their Asian or White peers. When considering reading achievement tests, White males and females essentially form a cluster of achievement above that for Hispanic and Black males and females. In fact, Black males show the lowest level of performance across 13 years of data (see Meade, 2006, Figure 3). Therefore, examining potential variations in the magnitude of gender differences in school achievement as a function of racial composition, at least for samples from the United States, is crucial to better inform critical areas of concerns for school performance. Based on Mead's report, it would not be surprising to find that the gender differences are largest among samples composed of a majority of Black students in the United States.

Other Potential Moderators

When we reviewed the literature on school achievement and gender, only a limited set of moderators directly relevant to gender differences became apparent. In fact, much of the research did not focus on gender differences. Accordingly, other potential moderators of general applicability were considered as exploratory variables.

A number of authors have suggested that socioeconomic status (SES) might relate to school achievement (e.g., Dewaele, 2007; Fischbein, 1990; Undheim & Nordvik, 1992). Specifically, the underlying reasoning is that values and beliefs held by parents with a higher SES might contribute to better school achievement in their children. However, SES is rarely reported directly in research and could not be used for all the studies sampled here. In contrast, school type (private or public) is always reported or can be determined with further research on the nature of the schools involved in data collection. Accordingly, this dichotomy provided an indirect measure of SES following the notion that, on average, students in a private school should typically come from a higher SES family than those in public schools. Based on this classification, the general expectation is that private school samples should achieve better grades than public school samples. From this perspective, it is possible that ceiling effects found in private school samples could potentially reduce the magnitude of gender differences. This possibility was explored here.

School achievement can be measured on a seemingly infinite number of scales. Specifically, some schools rely on percentage marks, others use a 4-point scale, while others report marks on a 12-point scale, and so on. Although this is a strictly statistical moderator, it has to be considered as the underlying variation in range of scores could potentially affect the magnitude of the effects. However, this scale of measurement moderator should be considered more of a control variable than a meaningful factor with cognitive implications.

Some research on gender stereotypes suggests that males and females tend to expect the gender composition (male to female ratio) to vary depending on specific areas of study. For example, S. Beyer (1999) reported that undergraduate participants estimated that gender composition would be about 61% females in an English major when the actual value was 64%. In contrast, for biology, the expected percentage of females was estimated at 42.2% by female participants and 45.1% by male participants when the actual value was 59.6%. In a sense, this perception that there are male- or female-dominated areas could reflect highly publicized findings of gender differences in these domains. However, the data reported by Beyer also indicated that there is some truth in the stereotypical expectations, at least at the university level. Therefore, considering the ratio of males to females in each sample might provide a further, indirect measure of whether an area is female or male dominated. Of course, the male to female ratio is bound to reflect the availability of volunteer participants in many of the studies. However, much of the retrieved data came from whole classes so that it would reflect class composition. In any case, a significant effect of this variable as a moderator would provide an indirect validity check. This variable also provides additional information to interpret the results. Accordingly, gender composition of the samples in terms of male to female ratio was considered as a potential moderator.

Current Meta-Analysis

The present analysis aimed to provide a summary of findings pertaining to gender differences in scholastic achievement. In doing so, we attempted to provide an exhaustive sample of the published literature providing relevant data. The quantification of these gender differences and the identification of relevant moderator variables formed the primary goals of the analysis.

One novel aspect of the work presented here lies in the fact that, to our knowledge, no such meta-analysis has been published to date as those that have been published focused on achievement tests. Therefore, in the present study, effect sizes were derived exclusively from the school marks obtained from teachers rather than from individual tests. As such, a new light is shed on gender differences in school achievement. It should be noted that this approach also limits the number of relevant studies as it eliminates those relying exclusively on self-reported grades. Essentially, issues relevant to social desirability would add a potential confound, limiting conclusions that can be drawn from the findings. Similarly, only typical samples were included as the inclusion of atypical samples such as remedial groups or gifted samples would also add extraneous variances. Finally, a numerical value was required for effect-size calculation (either directly available or converted from a letter grade). As numerical values are generally unavailable for preschool and kindergarten samples, data retrieval started with elementary school samples.

A second crucial novel aspect of the present analysis is its reliance on powerful approaches to meta-analysis. In particular, a direct comparison of course material has often been impossible. Specifically, in the research sampled here, effect sizes for different types of course material (e.g., math, science, and reading) were often obtained from the same samples. When using conventional meta-analytic techniques, this violation of the independence of effects assumption would invalidate the results (Borenstein, Hedges, Higgins, & Rothstein, 2009). Accordingly, in the overall sample, the present analysis applied hierarchical linear modeling to the meta-analysis (also known as multilevel meta-analysis; see Hox, 2010; Raudenbush & Bryk, 2002). This approach makes no assumption concerning independence of effects and is ideal for the analysis of meta-analytic data as they follow a clear hierarchical structure. This technique was necessary to examine variations in the magnitude of gender differences across course materials. In addition, mixed-effects meta-analysis was applied to examine the influence of moderator variables within course material as it was expected that effect sizes within these groupings would be independent. Therefore, the combination of approaches to meta-analysis provides a powerful analytic strategy that maximizes the amount of information gained from the research presented here.

Method

Study Selection

Retrieving studies initially involved searching for periodicals in the databases of PsycINFO, PsycARTICLES, and ERIC using the search terms *sex, gender, sex differences,* or *gender differences* with *school grades, GPA, school achievement,* or *school marks*. The initial search excluded theses, government reports, books, magazine articles, and any other clearly nonrefereed source. However, most government reports are represented in refereed sources (e.g., Balsa, Giuliano, & French, 2011; Catsambis, 1994; Keith & Benson, 1992) so that they were actually included in the present meta-analysis. These searches resulted in 6,048 nonoverlapping hits. However, as the inclusion of *sex, gender, sex differences,* or *gender differences* might have limited the hits to those where gender differences were central to the research question, a second set of searches was performed without these terms. An additional 8,994 hits resulted from this new search set. The searches included foreign-language articles as the databases provided an English abstract for all of them. In addition, theses and dissertations were considered as a possible source of unpublished material. Therefore, altogether, 15,042 published articles (including 753 articles in languages other than English) and 2,265 theses and dissertations were reviewed for possible inclusion in the meta-analysis.

In addition, researchers whose articles were retrieved for inclusion in the analysis and who published this work in the last 10 years were contacted by e-mail with a request for similar unpublished research. Furthermore, all researchers contacted for other reasons (clarifications, additional information on data reported, etc.) received a similar request. Altogether, 118 researchers were contacted directly. A posting requesting unpublished research was also sent to the following electronic mailing lists: American Education Research Association, Spatial Learning Network, Bilingualism and Bilingual Education Network, Educational Research List, Athens Institute for Education and Research, European Early Childhood Education Research Association (plus the Hong Kong and Japan chapters of this association), and the Alliance for International Education. Finally, the first author of the present article requested unpublished research of relevance as a question topic on his Research Gate webpage. As a result of these efforts, we received only eight responses to our request, and of those, only one provided previously unpublished data (three effect sizes). Thankfully, 25 unpublished dissertations relevant to our purpose were retrieved in the electronic search. Therefore, the final data set included a small subset of unpublished research.

Specific selection criteria determined whether a study could be included in the final sample in order to control for extraneous variables and ensure validity. Accordingly, one of the two authors carefully read the abstract for each study as a first step in determining if the inclusion criteria were met. When fit with the inclusion criteria was unclear from the abstract, the actual article was consulted.

The specific criteria used in making inclusion decisions required studies to have both male and female participants in Grade 1 (elementary school) or later. A study had to report teacher-assigned official subject/school marks or global GPA to be included. One-time measures such as Scholastic Aptitude Test scores and national assessment tests were therefore excluded. In addition, the use of official marks excluded studies that relied on self-reported values.

As is always the case with a meta-analysis, the studies had to report usable data for calculating the effect size (see the Measure of Effect Size section) to be included. For example, some studies dichotomized GPA as low or high, whereas others did not report the direction of the effect. When information was missing and the study was published in 2000 or more recently, the first author was contacted. A publication year of 2000 was deemed recent enough so that the data might still be available to the authors. Only seven authors had to be contacted in this manner, and five of them provided clarifications that resulted in usable data.

Specific exclusion criteria were also defined. In particular, studies that reported on special populations were not included. For example, some studies had a selection criterion such as high-risk or mentored students or students born at low birth weights. Such studies were excluded so that the present results can be interpreted as reflecting what is found in the average, typical student. It should be noted, however, that if a study reported on a control group that met the other criteria, data from such a group were included. When a study reported on a longitudinal sample, only the first year of data collection was used. This decision was based on the notion that aging effects are clearly beyond the scope of the present report and extensive inclusion of longitudinal data would introduce extraneous variance. Similarly, when multiple articles reported on the same sample, only the first study reporting on these data was included to avoid data duplication.

Reference lists obtained from the initial search were also used to retrieve a number of relevant studies. The data-collection window ended in August 2011 and, with the application of the inclusion/exclusion criteria, resulted in a final sample of 502 effect sizes drawn from 369 samples. The actual effect sizes included in the final sample are presented in Table 1, along with the moderators that produced significant findings. Out of the 502 effect sizes, 33 came from unpublished research in English (30 from dissertations, three from one unpublished paper). For the remainder, 436 effect sizes were from papers published in English, and 33 were from work published in other languages (Chinese, French, German, and Spanish).

Coding of Variables

A number of sample-level and measure-level variables were coded to assist in the goal of identifying factors that might moderate gender differences in school achievement. Sample-level variables would be characteristics inherent in the samples themselves, such as mean age, race/ethnicity, and so on. Similarly, measure-level variables reflect factors that are inherent in the school marks themselves, such a scale of measurement, their source, the content area, and so on.

Sample-level variables. Considering the variety of research questions investigated in the retrieved literature, the actual national origin of the participants was often unavailable. However, the country where testing took place was always mentioned, if only in the first author's affiliation. Accordingly, when nationality was not explicitly reported, the country of testing was used in lieu of national origin, based on the rationale that the majority of the sample would originate in this country. However, when this variable was considered more closely, it turned out that 258 out of 369 samples (69.9%) originated in the United States. Other regions that were represented involved samples from Norway ($k = 21$), Canada (13), Turkey (eight), Germany (13), Taiwan (six), Malaysia (six), Israel (five), New Zealand or Australia (five), Sweden (five), Slovakia (four), United Kingdom (three), Africa (three), Finland (three), and multiple countries (two). The remaining countries were represented in only one sample (Belgium, Czech Republic, Estonia, Mexico, Hong Kong, India, Iran, Jordan, the Netherlands, Portugal, Saudi Arabia, Serbia, and Slovenia). Coding of all these countries would produce essentially uninterpretable results. Accordingly, nationality was coded into three categories. North America included the United States and Canada; Scandinavia was formed by grouping samples from Norway, Finland, and Sweden; and the remaining regions were grouped in what was essentially an "other countries excluding North America and Scandinavia" category. Although crude and admittedly post hoc, it was hoped that this categorization might distinguish between the reputedly more egalitarian Scandinavian class structure and educational systems (e.g., Nordvik & Amponsah, 1998; Wiborg, 2004) and other, presumably less egalitarian societies.

Year of publication was also coded routinely and was included as a potential moderator. This factor could potentially reflect social changes that might promote fluctuations in gender differences on school marks.

In coding racial composition of the sample, we followed an approach similar to that proposed by Goldberg, Prause, Lucas-Thompson, and Himsel (2008). Therefore, United States samples were coded as composed of participants who were 75% or more White, 75% or more Black, 75% or more Hispanic, 75% or more Asian, racially diverse (no majority), or not reported. Samples from countries other than the United States were coded as non-U.S. samples so that the whole sample of retrieved effect sizes could be considered in the analysis.

Although it might seem appealing to code gender composition as a continuous variable reflecting the male to female ratio, in some cases the number of males and females in a sample had to be estimated as equal as only a total sample size was provided. Such an estimation of sample size was only implemented to make it possible to calculate the weights required in meta-analysis. However, in the context of gender ratio as a moderator, such estimations could potentially introduce extraneous variance. Accordingly, gender composition was coded into four categories: more females than males, equal number of each gender, more males than females, and estimated composition of samples. Of further relevance to sample size and composition, it should be noted that when the sample size was given as a range, the lower value was used as a conservative measure.

A final sample characteristic considered whether a sample was obtained in a public or private/parochial school. In the rare cases when type of school was not clearly labeled as public or private, it was coded as public. As previously mentioned, this variable was used as an indirect measure of SES.

Measure-level variables. At the level of measure-specific characteristics, the course material or subject in which the mark was obtained was coded as a global measure (typically reflecting a GPA), language (including, e.g., marks obtained in native-language and foreign-language courses), mathematics (also including economics and statistics courses), science (including both specific courses in biology, physics, etc., as well as general science

Table 1
Studies Included in the Present Analysis

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Abbott (1981)	yes	54	62	North America	elementary	Black	f > m	global	.522
Abedi (1991)	yes	3,316	3,145	North America	graduate	diverse	m > f	global	.000
D. Adams, Astone, Nunez-Wormack, & Smodlaka (1994)	yes	237	253	North America	high school	Hispanic	f > m	global	.387
R. E. Adams & Laursen (2007)	yes	231	238	North America	middle/junior	diverse	f > m	global	.303
Ajiboye & Tella (2006)	yes	22	54	Other	undergraduate	non-U.S.	f > m	socsci	−.217
Alessandri (1992)	yes	72	72	North America	elementary	diverse	f = m	global	.971
Alfan & Othman (2005)	yes	66	248	Other	undergraduate	non-U.S.	f > m	global	.285
Al-Mannie (1990)	yes	15,120	5,354	Other	undergraduate	non-U.S.	m > f	global	.064
Alnabhan, Al-Zegoul, & Harwell (2001)	yes	300	300	Other	undergraduate	non-U.S.	f > m	global	.000
Altschul, Oyserman, & Bybee (2006)	yes	69	70	North America	middle/junior	diverse	f > m	global	.571
Anderson (2006)	yes	1,610	1,487	North America	graduate	NR	m > f	global	.335
Annor (2010)	no	55	35	North America	undergraduate	diverse	m > f	global	.153
Ari, Atalay, & Aljamhan (2010)	yes	22	91	North America	undergraduate	NR	f > m	global	.330
Ari, Atalay, & Aljamhan (2010)	yes	22	91	North America	high school	NR	f > m	global	.359
Arrison (1998)	no	207	208	North America	undergraduate	NR	f = m	global	.697
Attaway & Bry (2004)	yes	31	28	North America	middle/junior	Black	m > f	global	.046
Aunio & Niemivirta (2010)	yes	105	107	Scandinavia	elementary	non-U.S.	f > m	math	.067
Ayers, Bustamante, & Campana (1973)	yes	112	112	North America	undergraduate	NR	estimated	lang	.539
Ayers & Quattlebaum (1992)	yes	60	7	North America	graduate	NR	m > f	global	.000
Azen, Bronner, & Gafni (2002)	yes	26,278	35,607	Other	undergraduate	non-U.S.	f > m	global	−.077
Babaoye (2000)	no	880	847	North America	undergraduate	Black	m > f	global	.342
Balsa, Giuliano, & French (2011)	yes	2,049	2,243	North America	high school	diverse	f > m	global	.314
Banducci (1967)	yes	1,520	1,494	North America	high school	NR	m > f	global	.309
Banreti-Fuchs (1972)	yes	356	304	North America	high school	non-U.S.	m > f	global	.000
Baucal, Pavlovic-Babic, & Willms (2006)	yes	1,995	1,994	Other	middle/junior	non-U.S.	estimated	global	.062
Bean & Bradley (1986)	yes	571	947	North America	undergraduate	White	f > m	global	.096
Beaudin, Horvath, & Wright (1992)	yes	204	119	North America	undergraduate	White	m > f	math	−.358
Beecher & Fischer (1999)	yes	178	231	North America	high school	NR	f > m	global	.000
Beecher & Fischer (1999)	yes	178	231	North America	undergraduate	NR	f > m	global	.000
Beer (1989)	yes	24	28	North America	elementary	NR	f > m	global	1.186
Behrens & Vernon (1978)	yes	155	137	North America	middle/junior	non-U.S.	m > f	lang	.709
Behrens & Vernon (1978)	yes	155	137	North America	middle/junior	non-U.S.	m > f	math	.220
Benedict & Hoag (2004)	yes	116	82	North America	undergraduate	NR	m > f	math	−.242
Bennett, Gottesman, Rock, & Cerullo (1993)	yes	84	86	North America	elementary	diverse	f > m	global	.154
Bennett, Gottesman, Rock, & Cerullo (1993)	yes	89	93	North America	elementary	diverse	f > m	global	.400
Bennett, Gottesman, Rock, & Cerullo (1993)	yes	88	78	North America	elementary	diverse	m > f	global	.286
Bennett, Gottesman, Rock, & Cerullo (1993)	yes	45	37	North America	elementary	diverse	m > f	global	−.125
Bennett, Gottesman, Rock, & Cerullo (1993)	yes	60	53	North America	elementary	diverse	m > f	global	.180
Bennett, Gottesman, Rock, & Cerullo (1993)	yes	43	38	North America	elementary	diverse	m > f	global	.200
Berndt & Miller (1990)	yes	54	99	North America	middle/junior	White	f > m	lang	.345
Betts & Morell (1999)	yes	2,738	2,885	North America	undergraduate	diverse	f > m	global	.128
H. N. Beyer (1971)	yes	695	756	North America	undergraduate	NR	f > m	global	.170
Bienstock, Martin, Tzou, & Fox (2002)	yes	193	62	North America	graduate	NR	m > f	other/NR	.480
Bink, Biner, Huffman, Geer, & Dean (1995)	yes	37	69	North America	undergraduate	NR	f > m	global	.181
Blackman, Hall, & Darmawan (2007)	yes	25	154	Other	undergraduate	non-U.S.	f > m	global	.000
Blair & Millea (2004)	yes	2,990	2,516	North America	undergraduate	White	m > f	global	.559
Blaser (1981)	yes	72	71	North America	undergraduate	NR	estimated	global	.479
Boldt (2000)	no	79	141	North America	undergraduate	diverse	f > m	global	.039
Booth (1983)	yes	25	23	North America	high school	NR	m > f	global	.205
Borde (1998)	yes	185	164	North America	undergraduate	NR	m > f	other/NR	.075
Borup (1971)	yes	260	260	North America	undergraduate	diverse	estimated	global	.445
Bourquin (1999)	no	60	170	North America	undergraduate	NR	f > m	math	.274
Bowman & Partin (1993)	yes	40	40	North America	undergraduate	NR	f = m	global	.456

(*table continues*)

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Brady, Tucker, Harris, & Tribble (1992)	yes	331	349	North America	middle/junior	diverse	f > m	global	.000
Bridgeman & Lewis (1994)	yes	208	273	North America	undergraduate	NR	f > m	lang	.040
Bridgeman & Lewis (1994)	yes	131	169	North America	undergraduate	NR	f > m	science	−.030
Bridgeman & Lewis (1994)	yes	896	677	North America	undergraduate	NR	m > f	socsci	.093
Bridgeman & Wendler (1991)	yes	1,678	1,704	North America	graduate	NR	f > m	math	.115
Britner (2008)	yes	233	269	North America	high school	White	f > m	science	.309
Brog (1985)	yes	32	32	North America	middle/junior	NR	f = m	global	.504
Brog (1985)	yes	32	32	North America	high school	NR	f = m	global	.087
Brooks (1987)	yes	154	168	North America	undergraduate	NR	f > m	math	.591
Brooks & Mercincavage (1991)	yes	20	40	North America	undergraduate	NR	f > m	lang	.000
Brooks & Mercincavage (1991)	yes	206	199	North America	undergraduate	NR	m > f	math	.433
Brooks & Mercincavage (1991)	yes	120	149	North America	undergraduate	NR	f > m	math	.267
Brooks & Rebeta (1991)	yes	312	283	North America	undergraduate	NR	m > f	socsci	.551
Brown & Jones (2004)	yes	152	182	North America	high school	Black	f > m	global	.324
Brubeck & Beer (1992)	yes	13	18	North America	high school	White	f > m	global	.668
Brubeck & Beer (1992)	yes	12	18	North America	high school	White	f > m	global	.476
Brubeck & Beer (1992)	yes	17	18	North America	high school	White	f > m	global	1.011
Brubeck & Beer (1992)	yes	13	17	North America	high school	White	f > m	global	.104
Buck (1985)	yes	209	256	North America	undergraduate	NR	f > m	math	.000
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	lang	.205
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	lang	.122
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	lang	.205
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	lang	.205
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	math	.122
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	science	.000
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	socsci	.000
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	other/NR	.000
Bunce & Calvert (1974)	yes	594	461	Other	middle/junior	non-U.S.	m > f	other/NR	.205
Burgert (1935)	yes	95	96	North America	middle/junior	NR	f > m	lang	.789
Burgert (1935)	yes	95	96	North America	middle/junior	NR	f > m	math	.412
Burgert (1935)	yes	95	96	North America	middle/junior	NR	f > m	science	.361
Burgert (1935)	yes	95	96	North America	middle/junior	NR	f > m	socsci	.294
Burgert (1935)	yes	95	96	North America	middle/junior	NR	f > m	other/NR	.667
Burgert (1935)	yes	95	96	North America	middle/junior	NR	f > m	other/NR	.750
Burgette & Magun-Jackson (2008)	yes	477	672	North America	undergraduate	diverse	f > m	global	.116
Burke (1989)	yes	585	660	North America	middle/junior	NR	f > m	lang	.592
Burke (1989)	yes	278	313	North America	middle/junior	NR	f > m	lang	.599
Burke (1989)	yes	585	660	North America	middle/junior	NR	f > m	math	.230
Burke (1989)	yes	585	660	North America	middle/junior	NR	f > m	science	.356
Burke (1989)	yes	585	660	North America	middle/junior	NR	f > m	socsci	.400
Bursik & Martin (2006)	yes	64	78	North America	high school	White	f > m	global	.372
Burts et al. (1993)	yes	79	87	North America	elementary	diverse	f > m	lang	.602
Burts et al. (1993)	yes	79	87	North America	elementary	diverse	f > m	lang	.545
Burts et al. (1993)	yes	79	87	North America	elementary	diverse	f > m	lang	.358
Burts et al. (1993)	yes	79	87	North America	elementary	diverse	f > m	math	.372
Burts et al. (1993)	yes	79	87	North America	elementary	diverse	f > m	science	.421
Burts et al. (1993)	yes	79	87	North America	elementary	diverse	f > m	socsci	.472
Buseman & Harders (1932)	yes	840	787	Other	elementary	non-U.S.	m > f	global	.302
Büyüköztürk (2004)	yes	107	141	Other	high school	non-U.S.	f > m	global	.747
Büyüköztürk (2004)	yes	139	163	Other	high school	non-U.S.	f > m	global	.676
Byrd & Chavous (2009)	yes	315	248	North America	middle/junior	Black	m > f	global	.606
Byrns (1930)	yes	27,854	27,854	North America	undergraduate	NR	estimated	global	.017
Calafiore & Damianov (2011)	yes	180	258	North America	undergraduate	Hispanic	f > m	math	−.109
Call, Beer, & Beer (1994)	yes	63	52	North America	elementary	NR	m > f	global	.793
Cantwell, Archer, & Bourke (2001)	yes	3,515	4,785	Other	undergraduate	non-U.S.	f > m	global	.254
Caso-Niebla & Hernández-Guzmán (2007)	yes	681	792	Other	high school	non-U.S.	f > m	global	.427
Catsambis (1994)	yes	8,146	7,951	North America	middle/junior	diverse	m > f	math	.068
Cauce (1987)	yes	42	47	North America	middle/junior	Black	f > m	global	.000
Chang & Chen (1977)	yes	192	180	Other	elementary	non-U.S.	m > f	lang	.129
Chang & Chen (1977)	yes	196	167	Other	elementary	non-U.S.	m > f	lang	.245
Chang & Chen (1977)	yes	178	203	Other	elementary	non-U.S.	f > m	lang	−.032
Chang & Chen (1977)	yes	205	181	Other	elementary	non-U.S.	m > f	lang	.057
Chang & Chen (1977)	yes	205	181	Other	elementary	non-U.S.	m > f	lang	.254
Chang & Chen (1977)	yes	179	147	Other	elementary	non-U.S.	m > f	lang	−.055

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Chang & Chen (1977)	yes	192	180	Other	elementary	non-U.S.	m > f	math	−.236
Chang & Chen (1977)	yes	196	167	Other	elementary	non-U.S.	m > f	math	−.231
Chang & Chen (1977)	yes	178	203	Other	elementary	non-U.S.	f > m	math	−.138
Chang & Chen (1977)	yes	205	181	Other	elementary	non-U.S.	m > f	math	−.305
Chang & Chen (1977)	yes	205	181	Other	elementary	non-U.S.	m > f	math	−.246
Chang & Chen (1977)	yes	179	147	Other	elementary	non-U.S.	m > f	math	−.183
J. A. Chen & Pajares (2010)	yes	297	211	North America	elementary	diverse	m > f	science	.061
M. Chen & Ehrenberg (1993)	yes	365	388	Other	elementary	non-U.S.	f > m	lang	.038
Cheung & McBride-Chang (2008)	yes	49	42	Other	elementary	non-U.S.	m > f	global	.178
Chittum (1996)	no	37	29	North America	undergraduate	NR	m > f	global	.620
Cicirelli (1977)	yes	80	80	North America	elementary	NR	f = m	global	.523
Coates & Southern (1972)	yes	198	166	North America	undergraduate	NR	m > f	lang	.000
Cogan (2010)	yes	1,035	1,035	North America	undergraduate	White	estimated	global	.391
Craddick (1966)	yes	60	60	North America	undergraduate	NR	f = m	global	.478
Cruickshank, Kennedy, & Kapel (1980)	yes	248	202	North America	undergraduate	NR	m > f	global	.278
Cruickshank, Kennedy, & Kapel (1980)	yes	74	76	North America	graduate	NR	f > m	global	.000
Daly (2009)	no	123	415	North America	undergraduate	White	f > m	global	.214
Dambrot, Silling, & Zook (1988)	yes	71	119	North America	high school	NR	f > m	global	.941
Dambrot, Silling, & Zook (1988)	yes	71	119	North America	undergraduate	NR	f > m	other/NR	.286
Daubman, Heatherington, & Ahn (1992)	yes	72	80	North America	undergraduate	Black	f > m	global	−.163
Daubman, Heatherington, & Ahn (1992)	yes	69	80	North America	undergraduate	Black	f > m	global	.163
Davidson & Haffey (1979)	yes	29	36	North America	high school	White	f > m	science	.380
Day (1999)	no	118	134	North America	undergraduate	White	f > m	global	.337
Daymont & Blau (2008)	yes	136	109	North America	undergraduate	NR	m > f	other/NR	.100
DeBerard, Spielmans, & Julka (2004)	yes	57	147	North America	high school	White	f > m	global	.283
DeBerard, Spielmans, & Julka (2004)	yes	57	147	North America	undergraduate	White	f > m	global	.516
DeCoster (1979)	yes	111	103	North America	undergraduate	NR	m > f	global	−.204
Demirbas & Demirkan (2007)	yes	58	53	Other	undergraduate	non-U.S.	m > f	global	.000
Demirbas & Demirkan (2007)	yes	51	37	Other	undergraduate	non-U.S.	m > f	global	.000
Demirbas & Demirkan (2007)	yes	24	50	Other	undergraduate	non-U.S.	f > m	global	.852
Deslandes, Bouchard, & St-Amant (1998)	yes	243	282	North America	high school	non-U.S.	f > m	lang	.626
Deslandes, Bouchard, & St-Amant (1998)	yes	243	282	North America	high school	non-U.S.	f > m	math	.107
Desler & North (1978)	yes	3,185	3,176	North America	undergraduate	NR	m > f	global	.000
Dewaele (2007)	yes	42	47	Other	high school	non-U.S.	f > m	lang	.467
Dewaele (2007)	yes	42	47	Other	high school	non-U.S.	f > m	lang	−.037
de Wolf (1981)	yes	953	1,122	North America	high school	NR	f > m	math	.088
Di Lorenzo (2009)	no	72	219	North America	undergraduate	diverse	f > m	global	.305
Ding (2008)	yes	153	145	North America	middle/junior	White	m > f	lang	.324
Ding (2008)	yes	129	134	North America	middle/junior	White	f > m	math	.220
Ding, Song, & Richardson (2006)	yes	234	224	North America	middle/junior	White	m > f	math	.315
Ding, Song, & Richardson (2006)	yes	157	210	North America	high school	White	f > m	math	.382
Dubey (1982)	yes	20	20	Other	undergraduate	non-U.S.	f = m	global	.806
Duckworth & Seligman (2006)	yes	62	78	North America	middle/junior	diverse	f > m	lang	.728
Duckworth & Seligman (2006)	yes	14	13	North America	middle/junior	diverse	m > f	math	.518
Duckworth & Seligman (2006)	yes	47	66	North America	middle/junior	diverse	f > m	math	.583
Duckworth & Seligman (2006)	yes	62	78	North America	middle/junior	diverse	f > m	socsci	.612
Duckworth & Seligman (2006)	yes	75	89	North America	middle/junior	diverse	f > m	lang	.723
Duckworth & Seligman (2006)	yes	59	74	North America	middle/junior	diverse	f > m	math	.336
Duckworth & Seligman (2006)	yes	16	15	North America	middle/junior	diverse	m > f	math	.667
Duckworth & Seligman (2006)	yes	75	89	North America	middle/junior	diverse	f > m	socsci	.478
Dunham (1973)	yes	161	142	North America	undergraduate	NR	m > f	global	.430
Dunkake, Kiechle, Klein, & Rosar (2012)	yes	38	39	Other	middle/junior	non-U.S.	f > m	global	.454
Edwards & Thacker (1979)	yes	195	137	North America	high school	NR	m > f	global	.294
Elliott & Strenta (1988)	yes	508	405	North America	undergraduate	NR	m > f	global	.083
Elmore & Vasu (1980)	yes	86	76	North America	graduate	NR	m > f	math	.515
Erkman, Caner, Sart, Börkan, & Şahan (2010)	yes	109	114	Other	elementary	non-U.S.	f > m	global	.160
Erkut (1983)	yes	176	116	North America	undergraduate	NR	m > f	global	.000
Farkas, Sheehan, & Grobe (1990)	yes	1,824	1,824	North America	middle/junior	NR	estimated	lang	.437

(*table continues*)

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Farkas, Sheehan, & Grobe (1990)	yes	3,004	3,005	North America	middle/junior	NR	f > m	lang	.547
Farkas, Sheehan, & Grobe (1990)	yes	562	562	North America	middle/junior	NR	estimated	math	.543
Farkas, Sheehan, & Grobe (1990)	yes	3,462	3,462	North America	middle/junior	NR	estimated	science	.404
Farkas, Sheehan, & Grobe (1990)	yes	3,740	3,739	North America	middle/junior	NR	estimated	socsci	.379
Farmer, Irwin, Thompson, Hutchins, & Leung (2006)	yes	142	250	North America	middle/junior	Black	f > m	global	.674
Fayowski & MacMillan (2008)	yes	644	615	North America	undergraduate	non-U.S.	m > f	math	.154
Feinberg & Halperin (1978)	yes	135	143	North America	undergraduate	NR	f > m	math	.366
Felson (1980)	yes	199	204	North America	middle/junior	White	f > m	global	.453
Fischbein (1990)	yes	523	462	Scandinavia	elementary	non-U.S.	m > f	lang	.276
Fischbein (1990)	yes	522	462	Scandinavia	elementary	non-U.S.	m > f	lang	−.177
Fischbein (1990)	yes	522	461	Scandinavia	elementary	non-U.S.	m > f	math	.335
Fischl & Sagy (2009)[a]	yes	36	171	Other	undergraduate	non-U.S.	f > m	global	1.393
Flexer (1984)	yes	61	63	North America	middle/junior	NR	f > m	math	.430
Frailey & Crain (1914)	yes	14	18	North America	high school	NR	f > m	lang	.211
Frailey & Crain (1914)	yes	14	18	North America	high school	NR	f > m	lang	.478
Frailey & Crain (1914)	yes	14	18	North America	high school	NR	f > m	math	.029
Frailey & Crain (1914)	yes	14	18	North America	high school	NR	f > m	math	.090
Frailey & Crain (1914)	yes	14	18	North America	high school	NR	f > m	socsci	−.024
Frenzel, Pekrun, & Goetz (2007)	yes	1,036	1,017	Other	elementary	non-U.S.	m > f	math	.000
Friedrichsen (1997)	no	102	134	North America	high school	diverse	f > m	global	.515
Friend (2009)	no	59	73	North America	elementary	Black	f > m	global	.776
Furnham, Chamorro-Premuzic, & McDougall (2002)	yes	23	70	Other	undergraduate	non-U.S.	f > m	global	.539
Gadzella, Cochran, Parham, & Fournet (1976)	yes	44	107	North America	undergraduate	NR	f > m	global	−.027
Gadzella, Williamson, & Ginther (1985)	yes	61	68	North America	undergraduate	NR	f > m	global	.449
Gerberich et al. (1997)	yes	118	45	North America	undergraduate	NR	m > f	global	.221
Giancarlo & Facione (2001)	yes	300	453	North America	undergraduate	NR	f > m	global	.129
Gifford, Briceno-Perriott, & Minzano (2006)	yes	1,448	1,618	North America	undergraduate	White	f > m	global	.071
Gillock & Reyes (1999)	yes	60	98	North America	high school	Hispanic	f > m	global	.396
Glass & Garrett (1995)	yes	81	91	North America	undergraduate	White	f > m	global	.224
Goetz, Frenzel, Hall, & Pekrun (2008)	yes	683	697	Other	middle/junior	non-U.S.	f > m	lang	.452
Goetz, Frenzel, Hall, & Pekrun (2008)	yes	683	697	Other	middle/junior	non-U.S.	f > m	math	−.128
A. Goodman & Koupil (2010)	yes	5,244	4,863	Scandinavia	elementary	non-U.S.	m > f	global	.066
S. B. Goodman & Cirka (2009)	yes	76	78	North America	undergraduate	diverse	f > m	lang	.000
Goodstein, Crites, & Heilbrun (1963)	yes	3,986	3,514	North America	undergraduate	NR	m > f	global	.252
Graham (1991)	yes	68	32	North America	graduate	White	m > f	global	.000
Grave (2011)	yes	6,439	4,858	Other	undergraduate	non-U.S.	m > f	global	−.037
Gupta, Harris, Carrier, & Caron (2006)	yes	271	180	North America	high school	NR	m > f	math	−.210
Ham (2004)	yes	93	106	North America	high school	diverse	f > m	global	.348
Hamre & Pianta (2001)	yes	91	88	North America	middle/junior	diverse	m > f	global	.348
Hancock (1999)	yes	149	120	North America	graduate	NR	m > f	global	.085
Hanna & Sonnenschein (1985)	yes	421	519	North America	middle/junior	NR	f > m	math	.147
Harris, Tanner, & Knouse (1996)	yes	182	216	North America	undergraduate	White	f > m	global	.307
Harrison (1996)	yes	139	140	North America	undergraduate	NR	f > m	other/NR	.000
Healy, Tullier, & Mourton (1990)	yes	222	304	North America	undergraduate	NR	f > m	global	.215
Heatherington et al. (1993)	yes	194	194	North America	undergraduate	White	f = m	global	.000
Heatherington et al. (1993)	yes	120	119	North America	undergraduate	White	m > f	global	.000
Heatherington, Townsend, & Burroughs (2001)	yes	40	46	North America	undergraduate	diverse	f > m	global	.497
Heaven, Ciarrochi, & Vialle (2007)	yes	382	394	Other	middle/junior	non-U.S.	f > m	lang	.387
Heaven, Ciarrochi, & Vialle (2007)	yes	382	394	Other	middle/junior	non-U.S.	f > m	math	.280
Heaven, Ciarrochi, & Vialle (2007)	yes	382	394	Other	middle/junior	non-U.S.	f > m	science	.086
Heaven, Ciarrochi, & Vialle (2007)	yes	382	394	Other	middle/junior	non-U.S.	f > m	socsci	.454
Heaven, Ciarrochi, & Vialle (2007)	yes	382	394	Other	middle/junior	non-U.S.	f > m	other/NR	.577
Helbig (2010)	yes	1,634	1,535	Other	elementary	non-U.S.	m > f	lang	.092
Helbig (2010)	yes	1,634	1,535	Other	elementary	non-U.S.	m > f	math	.000
Herron (1964)	yes	45	45	North America	undergraduate	NR	f = m	global	−.140
Hewitt & Goldman (1975)	yes	6,500	6,500	North America	undergraduate	NR	estimated	global	.073
Hildenbrand (2005)	no	769	769	North America	undergraduate	NR	estimated	global	.183
Hildenbrand (2005)	no	803	803	North America	undergraduate	NR	estimated	global	.342
Hildenbrand (2005)	no	808	807	North America	undergraduate	NR	estimated	global	.350
Hildenbrand (2005)	no	770	771	North America	undergraduate	NR	estimated	global	.386
Hildenbrand (2005)	no	821	820	North America	undergraduate	NR	estimated	global	.330

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Hogan et al. (2010)	yes	96	96	North America	high school	non-U.S.	f = m	global	.521
Horvath, Beaudin, & Wright (1992)	yes	265	159	North America	undergraduate	NR	m > f	math	−.251
Hosseini (1975)	yes	1,138	271	Other	high school	non-U.S.	m > f	global	.153
House & Keeley (1995)	yes	192	1,246	North America	graduate	NR	f > m	global	.456
Houston (1987)	yes	30	52	North America	undergraduate	White	f > m	global	.541
Hudy (2006)	no	696	1,036	North America	undergraduate	White	f > m	global	.080
Hughey (1995)	yes	88	130	North America	undergraduate	White	f > m	global	.317
Hunley et al. (2005)	yes	51	50	North America	high school	NR	estimated	global	.624
Huysamen & Roozendaal (1999)	yes	329	470	North America	undergraduate	White	f > m	global	.321
Imms (2000)	yes	793	1,438	Other	high school	non-U.S.	f > m	other/NR	.126
Ismail & Othman (2006)	yes	64	140	Other	undergraduate	non-U.S.	f > m	global	.000
Ismail & Othman (2006)	yes	74	187	Other	undergraduate	non-U.S.	f > m	global	.244
Ismail & Othman (2006)	yes	229	646	Other	undergraduate	non-U.S.	f > m	global	.133
Jacobowitz (1983)	yes	113	148	North America	middle/junior	Black	f > m	math	.148
Jacobowitz (1983)	yes	113	148	North America	middle/junior	Black	f > m	science	−.094
Jansen & Bruinsma (2005)	yes	78	218	Other	undergraduate	non-U.S.	f > m	global	.473
Johnson & Kuennen (2006)	yes	177	115	North America	undergraduate	White	m > f	math	.304
Johnston (1999)	no	31	191	North America	undergraduate	diverse	f > m	global	−.016
Jones (2010)	no	911	858	North America	high school	NR	m > f	global	.656
Kaczmarek & Franco (1986)	yes	18	25	North America	graduate	NR	f > m	global	−.203
Keiller (1997)	no	215	377	North America	undergraduate	White	f > m	global	.000
Keith & Benson (1992)	yes	6,392	6,760	North America	high school	White	f > m	global	.364
Keller, Crouse, & Trusheim (1993)	yes	1,633	1,632	North America	high school	NR	estimated	global	−.343
Keller, Crouse, & Trusheim (1993)	yes	1,633	1,632	North America	undergraduate	NR	estimated	global	.090
Kenney-Benson, Pomerantz, Ryan, & Patrick (2006)	yes	253	265	North America	elementary	White	f > m	math	.188
Khan, Haynes, Armstong, & Rohner (2010)	yes	169	182	North America	middle/junior	Black	f > m	lang	.546
Khan, Haynes, Armstong, & Rohner (2010)	yes	169	182	North America	middle/junior	Black	f > m	math	.478
Khan, Haynes, Armstong, & Rohner (2010)	yes	169	182	North America	middle/junior	Black	f > m	science	.367
King & Joshi (2008)	yes	620	120	North America	undergraduate	NR	m > f	science	.000
Kitsantas & Zimmerman (2009)	yes	56	167	North America	undergraduate	White	f > m	socsci	.054
Klugh & Bierley (1959)	yes	231	199	North America	undergraduate	NR	m > f	global	.595
Klugh & Bierley (1959)	yes	231	199	North America	high school	NR	m > f	global	.784
Koenig, Sireci, & Wiley (1998)	yes	630	479	North America	graduate	diverse	m > f	global	−.121
Koenig, Sireci, & Wiley (1998)	yes	630	479	North America	undergraduate	diverse	m > f	science	−.070
Koenig, Sireci, & Wiley (1998)	yes	6,637	4,642	North America	undergraduate	diverse	m > f	science	−.051
Kokkelenberg & Sinha (2010)	yes	20,261	23,784	North America	high school	diverse	f > m	global	.314
Kollárik (1991)	yes	52	56	Other	elementary	non-U.S.	f > m	lang	.369
Kollárik (1991)	yes	120	123	Other	elementary	non-U.S.	f > m	lang	.219
Kollárik (1991)	yes	52	56	Other	elementary	non-U.S.	f > m	math	.101
Kollárik (1991)	yes	120	123	Other	elementary	non-U.S.	f > m	math	.041
Kollárik (1991)	yes	52	56	Other	elementary	non-U.S.	f > m	science	.100
Kost, Pollock, & Finkelstein (2009)	yes	2,715	848	North America	undergraduate	White	m > f	science	−.115
Kucerova (1975)	yes	43	56	Other	elementary	non-U.S.	f > m	lang	.500
Kucerova (1975)	yes	60	55	Other	elementary	non-U.S.	m > f	lang	.248
Kucerova (1975)	yes	43	56	Other	elementary	non-U.S.	f > m	math	.125
Kucerova (1975)	yes	60	55	Other	elementary	non-U.S.	m > f	math	−.117
Kurdek & Sinclair (1988)	yes	96	123	North America	middle/junior	White	f > m	global	.080
Lagerström, Bremme, Eneroth, & Magnusson (1991)	yes	437	407	Scandinavia	elementary	non-U.S.	m > f	lang	.500
Lagerström, Bremme, Eneroth, & Magnusson (1991)	yes	437	407	Scandinavia	elementary	non-U.S.	m > f	math	.105
Lagerström, Bremme, Eneroth, & Magnusson (1991)	yes	437	407	Scandinavia	elementary	non-U.S.	m > f	socsci	.117
Lee & Nemzek (1941)	yes	150	150	North America	middle/junior	NR	f = m	lang	.779
Lee & Nemzek (1941)	yes	150	150	North America	middle/junior	NR	f = m	math	.248
Lee & Nemzek (1941)	yes	150	150	North America	middle/junior	NR	f = m	science	.398
Lee & Nemzek (1941)	yes	150	150	North America	middle/junior	NR	f = m	socsci	.378
Lehn, Vladovic, & Michael (1980)	yes	274	274	North America	high school	diverse	f = m	lang	.324
Lehn, Vladovic, & Michael (1980)	yes	274	274	North America	high school	diverse	f = m	math	.181
Lehn, Vladovic, & Michael (1980)	yes	274	274	North America	high school	diverse	f = m	socsci	.140

(*table continues*)

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Lehre, Hansen, & Laake (2009)	yes	6,075	16,155	Scandinavia	undergraduate	non-U.S.	f > m	global	.128
Lehre, Hansen, & Laake (2009)	yes	1,957	6,979	Scandinavia	undergraduate	non-U.S.	f > m	global	.237
Lehre, Hansen, & Laake (2009)	yes	23,931	39,794	Scandinavia	undergraduate	non-U.S.	f > m	global	−.107
Lehre, Hansen, & Laake (2009)	yes	23,591	42,768	Scandinavia	undergraduate	non-U.S.	f > m	global	−.027
Lehre, Hansen, & Laake (2009)	yes	62,799	39,349	Scandinavia	undergraduate	non-U.S.	m > f	global	.009
Lehre, Hansen, & Laake (2009)	yes	18,124	12,242	Scandinavia	undergraduate	non-U.S.	m > f	global	−.035
Lehre, Hansen, & Laake (2009)	yes	22,875	34,043	Scandinavia	undergraduate	non-U.S.	f > m	global	.010
Lehre, Hansen, & Laake (2009)	yes	17,131	30,239	Scandinavia	undergraduate	non-U.S.	f > m	global	.081
Lehre, Hansen, & Laake (2009)	yes	1,042	1,243	Scandinavia	undergraduate	non-U.S.	f > m	global	−.030
Lehre, Hansen, & Laake (2009)	yes	1,252	1,994	Scandinavia	undergraduate	non-U.S.	f > m	global	−.101
Lehre, Hansen, & Laake (2009)	yes	637	1,737	Scandinavia	graduate	non-U.S.	f > m	global	.003
Lehre, Hansen, & Laake (2009)	yes	791	3,600	Scandinavia	graduate	non-U.S.	f > m	global	−.022
Lehre, Hansen, & Laake (2009)	yes	4,731	7,622	Scandinavia	graduate	non-U.S.	f > m	global	−.172
Lehre, Hansen, & Laake (2009)	yes	3,609	6,528	Scandinavia	graduate	non-U.S.	f > m	global	−.037
Lehre, Hansen, & Laake (2009)	yes	11,964	6,461	Scandinavia	graduate	non-U.S.	m > f	global	−.108
Lehre, Hansen, & Laake (2009)	yes	7,830	4,121	Scandinavia	graduate	non-U.S.	m > f	global	.185
Lehre, Hansen, & Laake (2009)	yes	8,368	11,368	Scandinavia	graduate	non-U.S.	f > m	global	.007
Lehre, Hansen, & Laake (2009)	yes	4,713	5,941	Scandinavia	graduate	non-U.S.	f > m	global	.024
Lehre, Hansen, & Laake (2009)	yes	233	372	Scandinavia	graduate	non-U.S.	f > m	global	.243
Lehre, Hansen, & Laake (2009)	yes	474	651	Scandinavia	graduate	non-U.S.	f > m	global	.073
Lekholm & Cliffordson (2008)	yes	50,410	48,660	Scandinavia	high school	non-U.S.	m > f	lang	.503
Lekholm & Cliffordson (2008)	yes	50,410	48,660	Scandinavia	high school	non-U.S.	m > f	lang	.159
Lekholm & Cliffordson (2008)	yes	50,410	48,660	Scandinavia	high school	non-U.S.	m > f	math	.004
Lindley & Borgen (2002)	yes	104	209	North America	undergraduate	White	f > m	global	.409
Lindsay & Althouse (1969)	yes	226	88	North America	undergraduate	NR	m > f	global	.640
Lindsay & Althouse (1969)	yes	226	88	North America	high school	NR	m > f	global	.664
Llabre & Suarez (1985)	yes	72	112	North America	undergraduate	diverse	f > m	math	.343
Lloyd, Walsh, & Yailagh (2005)	yes	77	81	North America	middle/junior	non-U.S.	f > m	math	.442
Long, Monoi, Harper, Knoblauch, & Murphy (2007)	yes	123	132	North America	middle/junior	Black	f > m	global	.391
Long, Monoi, Harper, Knoblauch, & Murphy (2007)	yes	83	75	North America	high school	Black	m > f	global	.120
Lumme & Lehto (2002)	yes	33	33	Scandinavia	elementary	non-U.S.	f = m	lang	.731
Lumme & Lehto (2002)	yes	33	33	Scandinavia	elementary	non-U.S.	f = m	lang	.000
Lumme & Lehto (2002)	yes	33	33	Scandinavia	elementary	non-U.S.	f = m	math	−.643
Lumme & Lehto (2002)	yes	33	33	Scandinavia	elementary	non-U.S.	f = m	socsci	.000
Lunneborg (1977)	yes	898	735	North America	undergraduate	NR	m > f	global	.000
Lutz & Crist (2009)	yes	523	487	North America	high school	Hispanic	m > f	global	.208
Maqsud (1993)	yes	60	60	Other	middle/junior	non-U.S.	f = m	lang	−.166
Matthews (1991)	yes	376	420	North America	undergraduate	diverse	f > m	global	.196
Mau & Lynn (2001)	yes	4,256	5,494	North America	undergraduate	NR	f > m	global	.299
McCandless, Roberts, & Starnes (1972)	yes	221	222	North America	middle/junior	diverse	f > m	global	.500
McCornack & McLeod (1988)	yes	5,388	5,765	North America	undergraduate	NR	f > m	global	.062
McCornack & McLeod (1988)[a]	yes	5,388	5,765	North America	high school	NR	f > m	global	1.765
McDonald & McPherson (1975)	yes	122	30	North America	undergraduate	White	m > f	global	.387
Mickelson & Greene (2006)	yes	277	358	North America	middle/junior	Black	f > m	global	.453
Miller, Finley, & McKinley (1990)	yes	465	650	North America	undergraduate	NR	f > m	global	.000
Mills, Heyworth, Rosenwax, Carr, & Rosenberg (2009)	yes	102	279	Other	undergraduate	non-U.S.	f > m	global	.265
Morganson, Jones, & Major (2010)	yes	634	157	North America	undergraduate	diverse	m > f	global	−.040
Mpofu, D'Amico, & Cleghorn (1996)	yes	156	131	Other	undergraduate	non-U.S.	m > f	global	.447
Mullola et al. (2011)	yes	204	220	Scandinavia	high school	non-U.S.	f > m	lang	.742
Mullola et al. (2011)	yes	324	306	Scandinavia	high school	non-U.S.	m > f	math	−.015
Neighbors, Forehand, & Armistead (1992)	yes	25	33	North America	elementary	NR	f > m	global	.635
Nelson (1969)	yes	156	156	North America	high school	White	f = m	global	.182
Nguyen, Allen, & Fraccastoro (2005)	yes	98	102	North America	undergraduate	diverse	f > m	other/NR	.387
Nicpon et al. (2006)	yes	112	192	North America	undergraduate	White	f > m	global	.164
Nishina, Juvonen, & Witkow (2005)	yes	687	839	North America	elementary	diverse	f > m	global	.413
Odell & Schumacher (1998)	yes	140	124	North America	undergraduate	NR	m > f	math	.021
Öhlund & Ericsson (1994)	yes	299	299	Scandinavia	elementary	non-U.S.	estimated	global	.262
Olds & Shaver (1980)	yes	76	109	North America	undergraduate	NR	f > m	global	.286
O'Reilly & McNamara (2007)	yes	772	826	North America	high school	diverse	f > m	science	.228
Paolillo (1982)	yes	110	110	North America	graduate	NR	estimated	global	.283
Payne, Rapley, & Wells (1973)	yes	931	818	North America	undergraduate	NR	m > f	global	.338
Payne, Rapley, & Wells (1973)	yes	931	818	North America	high school	NR	m > f	global	.517

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Pedrini & Pedrini (1978)	yes	71	72	North America	undergraduate	diverse	f > m	global	.000
Perrault (1976)	no	118	128	North America	middle/junior	NR	f > m	global	.387
Phillips (1962)	yes	365	394	North America	middle/junior	NR	f > m	lang	.554
Phillips (1962)	yes	365	394	North America	middle/junior	NR	f > m	math	.247
Phillips (1962)	yes	365	394	North America	middle/junior	NR	f > m	socsci	.235
Pomerantz, Altermatt, & Saxon (2002)	yes	466	466	North America	elementary	White	f = m	lang	.280
Pomerantz, Altermatt, & Saxon (2002)	yes	466	466	North America	elementary	White	f = m	math	.138
Pomerantz Altermatt, & Saxon (2002)	yes	466	466	North America	elementary	White	f = m	science	.154
Pomerantz Altermatt, & Saxon (2002)	yes	466	466	North America	elementary	White	f = m	socsci	.164
Post et al. (2010)	yes	2,189	1,952	North America	undergraduate	White	m > f	math	.000
Preckel, Goetz, Pekrun, & Kleine (2008)	yes	82	78	Other	elementary	non-U.S.	m > f	lang	.657
Preckel, Goetz, Pekrun, & Kleine (2008)	yes	90	85	Other	elementary	non-U.S.	m > f	lang	−.542
Preckel, Goetz, Pekrun, & Kleine (2008)	yes	82	78	Other	elementary	non-U.S.	m > f	math	.136
Preckel, Goetz, Pekrun, & Kleine (2008)	yes	90	85	Other	elementary	non-U.S.	m > f	math	.014
Preiss & Fráňová (2006)	yes	304	331	Other	elementary	non-U.S.	f > m	global	.078
Preszler (2009)	yes	1,193	1,716	North America	undergraduate	diverse	f > m	science	.123
Pulvino & Hansen (1972)	yes	152	148	North America	high school	NR	m > f	global	.414
Quirk, Keith, & Quirk (2001)	yes	7,885	7,667	North America	high school	diverse	m > f	lang	.423
Quirk, Keith, & Quirk (2001)	yes	7,885	7,667	North America	high school	diverse	m > f	socsci	.211
Ramsbottom-Lucier, Johnson, & Elam (1995)	yes	373	184	North America	undergraduate	White	m > f	global	.581
Rech (1996)	yes	1,134	1,261	North America	undergraduate	NR	f > m	math	.101
Rech (1996)	yes	871	713	North America	undergraduate	NR	m > f	math	.036
Ritchie (2003)	no	118	196	North America	undergraduate	non-U.S.	f > m	global	.411
Rochelle & Dotterweich (2007)	yes	57	33	North America	undergraduate	White	m > f	math	−.197
Rogers, Theule, Ryan, Adams, & Keating (2009)	yes	110	121	North America	elementary	non-U.S.	f > m	lang	.606
Rogers, Theule, Ryan, Adams, & Keating (2009)	yes	110	121	North America	elementary	non-U.S.	f > m	math	.080
Rogers, Theule, Ryan, Adams, & Keating (2009)	yes	110	121	North America	elementary	non-U.S.	f > m	science	.366
Romine & Quattlebaum (1976)	yes	40	157	North America	high school	NR	f > m	global	.184
Romine & Quattlebaum (1976)	yes	40	157	North America	undergraduate	NR	f > m	global	.314
Romine & Quattlebaum (1976)	yes	48	77	North America	undergraduate	NR	f > m	global	.659
Romine & Quattlebaum (1976)	yes	48	77	North America	high school	NR	f > m	global	.939
Ross & Horner (1949)	yes	288	349	North America	undergraduate	NR	f > m	global	.156
Rothstein (2007)	yes	2,068	2,221	North America	high school	NR	f > m	global	.417
Ruban & McCoach (2005)	yes	119	256	North America	undergraduate	White	f > m	global	.432
Rustemeyer & Fischer (2005)	yes	89	86	Other	elementary	non-U.S.	f = m	math	.123
Rustemeyer & Fischer (2005)	yes	51	51	Other	middle/junior	non-U.S.	f = m	math	.155
Rustemeyer & Fischer (2005)	yes	32	35	Other	middle/junior	non-U.S.	f = m	math	.119
Sahin (2009)	yes	118	46	Other	undergraduate	non-U.S.	m > f	science	.408
Sahin (2009)	yes	70	30	Other	undergraduate	non-U.S.	m > f	science	.000
Sampson & Boyer (2001)	yes	57	103	North America	graduate	Black	f > m	global	.000
Saunders, Davis, Williams, & Williams (2004)	yes	74	95	North America	high school	Black	f > m	global	.431
Schaffer, Ahmadi, & Calkins (1986)	yes	203	173	North America	undergraduate	NR	m > f	global	.460
Schneider & Overton (1983)	yes	254	282	North America	undergraduate	NR	f > m	global	.129
Schulenberg, Asp, & Petersen (1984)	yes	85	103	North America	elementary	White	f > m	lang	.449
Schulenberg, Asp, & Petersen (1984)	yes	68	79	North America	elementary	White	f > m	lang	.276
Schulenberg, Asp, & Petersen (1984)	yes	85	103	North America	elementary	White	f > m	math	.093
Schulenberg, Asp, & Petersen (1984)	yes	68	79	North America	elementary	White	f > m	math	.147
Schulenberg, Asp, & Petersen (1984)	yes	85	103	North America	elementary	White	f > m	science	.095
Schulenberg, Asp, & Petersen (1984)	yes	68	79	North America	elementary	White	f > m	science	−.043
Schulenberg, Asp, & Petersen (1984)	yes	85	103	North America	elementary	White	f > m	socsci	.256
Schulenberg, Asp, & Petersen (1984)	yes	68	79	North America	elementary	White	f > m	socsci	.108
Scott (2010)	yes	32	42	North America	graduate	NR	f > m	global	.161
Seginer & Vermulst (2002)	yes	333	353	Other	middle/junior	non-U.S.	f > m	lang	.000
Sendelbach (1975)	yes	169	189	Other	elementary	non-U.S.	f > m	lang	−.145
Sendelbach (1975)	yes	169	189	Other	elementary	non-U.S.	f > m	lang	−.169

(*table continues*)

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Sendelbach (1975)	yes	169	189	Other	elementary	non-U.S.	f > m	lang	−.245
Sendelbach (1975)	yes	169	189	Other	elementary	non-U.S.	f > m	math	.022
Sendelbach (1975)	yes	169	189	Other	elementary	non-U.S.	f > m	science	−.129
Senfeld (1995)	no	92	159	North America	undergraduate	diverse	f > m	global	.349
Sexton & Goldman (1975)	yes	242	187	North America	high school	NR	m > f	lang	.529
Sexton & Goldman (1975)	yes	242	187	North America	high school	NR	m > f	lang	.425
Sexton & Goldman (1975)	yes	242	187	North America	high school	NR	m > f	math	.000
Sexton & Goldman (1975)	yes	242	187	North America	high school	NR	m > f	science	.139
Sexton & Goldman (1975)	yes	242	187	North America	high school	NR	m > f	socsci	.332
Seyfried (1998)	yes	57	56	North America	elementary	Black	m > f	global	.451
Sheard (2009)	yes	78	56	Other	undergraduate	non-U.S.	m > f	global	.377
J. A. Sherman (1980)	yes	115	140	North America	middle/junior	White	f > m	math	.075
L. W. Sherman & Hofmann (1980)	yes	92	82	North America	middle/junior	diverse	m > f	global	.366
Shields (2001)	yes	149	181	North America	undergraduate	White	f > m	global	.323
Shores, Smith, & Jarrell (2009)	yes	319	442	North America	elementary	diverse	f > m	math	−.326
Sibulkin & Butler (2008)	yes	42	166	North America	undergraduate	Black	f > m	math	.435
Simon (1978)	yes	41	45	North America	graduate	NR	f > m	global	.473
Simpkins, Davis-Kean, & Eccles (2006)	yes	104	123	North America	elementary	White	f > m	math	.000
Simpkins, Davis-Kean, & Eccles (2006)	yes	104	123	North America	elementary	White	f > m	science	−.065
Singleton (2007)	yes	320	360	North America	undergraduate	White	f > m	global	.451
Smith (2008)	no	240	162	North America	graduate	diverse	m > f	global	.164
Smrtnik-Vitulić & Zupančič (2011)	yes	82	127	Other	middle/junior	non-U.S.	f > m	global	.715
Snyder (2000)	yes	64	64	North America	high school	NR	estimated	global	.352
Soares, Guisande, Almeida, & Páramo (2009)	yes	140	305	Other	undergraduate	non-U.S.	f > m	global	.224
Steinmayr & Spinath (2008)	yes	138	204	Other	high school	non-U.S.	f > m	lang	.283
Steinmayr & Spinath (2008)	yes	138	204	Other	high school	non-U.S.	f > m	math	−.191
Stephens & Schaben (2002)	yes	68	68	North America	middle/junior	NR	f = m	global	.180
Stockard, Lang, & Wood (1985)	yes	106	85	North America	middle/junior	White	m > f	lang	.610
Stockard, Lang, & Wood (1985)	yes	138	121	North America	high school	White	m > f	math	.358
Stupnisky et al. (2007)	yes	304	498	North America	undergraduate	NR	f > m	global	.140
Stupnisky et al. (2007)	yes	304	498	North America	high school	NR	f > m	global	.242
Sulaiman & Mohezar (2006)	yes	253	236	Other	undergraduate	non-U.S.	m > f	global	.000
Sulaiman & Mohezar (2006)	yes	253	236	Other	graduate	non-U.S.	m > f	global	.000
Sullivan & Voyer (1998)	no	40	78	North America	high school	non-U.S.	f > m	lang	.463
Sullivan & Voyer (1998)	no	40	78	North America	high school	non-U.S.	f > m	math	.171
Sullivan & Voyer (1998)	no	22	35	North America	high school	non-U.S.	f > m	socsci	.804
Sullivan-Ham (2010)	no	162	302	North America	undergraduate	diverse	f > m	global	.000
Sweet & Nuttall (1971)	yes	206	180	North America	high school	NR	m > f	lang	.252
Talento-Miller (2008)	yes	832	231	Other	undergraduate	non-U.S.	m > f	global	−.232
Taube & Taube (1990)	yes	56	71	North America	undergraduate	diverse	f > m	global	.497
Taylor, Clay, Bramoweth, Sethi, & Roane (2011)	yes	217	621	North America	undergraduate	diverse	f > m	global	.292
Tenenbaum & Leaper (2003)	yes	16	17	North America	middle/junior	White	f > m	science	−.014
Thiede (1950)	yes	473	267	North America	undergraduate	NR	m > f	global	−.032
Thomas (1979)	yes	140	140	North America	undergraduate	Black	f = m	global	.784
Thompson, Samiratedu, & Rafter (1993)	yes	2,518	2,896	North America	undergraduate	diverse	f > m	global	.090
Ting & Robinson (1998)	yes	1,508	1,090	North America	high school	White	m > f	global	.086
Ting & Robinson (1998)	yes	1,375	1,007	North America	undergraduate	White	m > f	global	.127
Tiruneh (2007)	yes	476	475	North America	undergraduate	NR	estimated	socsci	.140
Treiber (2010)	no	39	29	North America	high school	NR	m > f	lang	1.105
Treiber (2010)	no	39	29	North America	high school	NR	m > f	math	.659
Trent (1974)	no	44	52	North America	high school	White	f > m	global	.640
Trippi & Baker (1989)	yes	117	193	North America	undergraduate	Black	f > m	global	.044
Trippi & Baker (1989)	yes	117	193	North America	high school	Black	f > m	global	.119
Truell, Zhao, Alexander, & Hill (2006)	yes	123	56	North America	graduate	NR	m > f	global	.000
Tulviste & Rohner (2010)	yes	109	115	Other	elementary	non-U.S.	f > m	global	.568
Undheim & Nordvik (1992)	yes	832	832	Scandinavia	middle/junior	non-U.S.	f = m	lang	.399
Undheim & Nordvik (1992)	yes	832	832	Scandinavia	middle/junior	non-U.S.	f = m	lang	.117
Undheim & Nordvik (1992)	yes	832	832	Scandinavia	middle/junior	non-U.S.	f = m	math	−.117
Urdan (1997)	yes	131	129	North America	middle/junior	White	m > f	global	.232
Valenzuela (1993)	yes	32	61	North America	middle/junior	Hispanic	f > m	lang	1.022
Véronneau & Dishion (2011)	yes	580	698	North America	elementary	White	f > m	global	.272
Violato (1990)	yes	836	3,651	North America	undergraduate	non-U.S.	f > m	global	.077
von Wittich (1972)	yes	592	303	North America	undergraduate	NR	m > f	lang	.372

Table 1 (*continued*)

Study	Pub	*N*m	*N*f	Nationality	Source	Racial composition	Gender composition	Course	*d*
Wang, Tu, & Shieh (2007)	yes	585	712	North America	undergraduate	NR	f > m	math	.310
Wang-Cheng, Fulkerson, Barnas, & Lawrence (1995)	yes	233	142	North America	graduate	NR	m > f	other/NR	.143
Warren, Jackson, & Sifers (2009)	yes	41	62	North America	middle/junior	diverse	f > m	global	.514
Wentzel (1991)	yes	220	203	North America	middle/junior	diverse	m > f	global	.345
Whalen-Schmeller (2006)	no	40	91	North America	undergraduate	Black	f > m	lang	.439
Whitley, Rawana, Pye, & Brownlee (2010)	yes	26	28	North America	elementary	non-U.S.	f > m	global	.913
Williams, Davis, Cribbs, Saunders, & Williams (2002)	yes	103	128	North America	high school	Black	f > m	global	.732
Witkow (2009)	yes	350	352	North America	high school	diverse	f > m	global	.182
Witt, Dunbar, & Hoover (1994)	yes	582	606	North America	high school	NR	f > m	lang	.321
Witt, Dunbar, & Hoover (1994)	yes	363	386	North America	high school	NR	f > m	math	−.007
Witt, Dunbar, & Hoover (1994)	yes	590	592	North America	high school	NR	f > m	science	−.076
Witt, Dunbar, & Hoover (1994)	yes	332	329	North America	high school	NR	m > f	socsci	−.044
Woodfield, Jessop, & McMillan (2006)	yes	308	323	Other	undergraduate	non-U.S.	f > m	global	.239
Woosley (2005)	yes	1,232	1,717	North America	undergraduate	White	f > m	global	.303
C. R. Wright & Houck (1995)	yes	18	20	North America	high school	NR	f > m	lang	.625
C. R. Wright & Houck (1995)	yes	17	19	North America	high school	NR	f > m	lang	1.023
C. R. Wright & Houck (1995)	yes	63	85	North America	high school	NR	f > m	lang	.973
C. R. Wright & Houck (1995)	yes	18	20	North America	high school	NR	f > m	math	.761
C. R. Wright & Houck (1995)	yes	17	19	North America	high school	NR	f > m	math	.493
C. R. Wright & Houck (1995)	yes	63	85	North America	high school	NR	f > m	math	.269
R. E. Wright & Palmer (1998)	yes	99	99	North America	graduate	NR	estimated	global	−.280
R. J. Wright & Bean (1974)	yes	884	747	North America	undergraduate	White	m > f	global	.386
Wynne (2003)	no	365	471	North America	undergraduate	diverse	f > m	global	.057
Yang & Lu (2001)	yes	292	103	North America	graduate	NR	m > f	global	.000
I. P. Young & Young (2010)	yes	45	55	North America	graduate	diverse	f > m	global	.369
I. P. Young & Young (2010)	yes	45	55	North America	undergraduate	diverse	f > m	global	.509
J. W. Young (1991)	yes	874	648	North America	high school	NR	m > f	global	.066
Zarb (1981)	yes	30	98	North America	high school	non-U.S.	f > m	global	.244
Zeidner (1987)	yes	505	683	Other	undergraduate	non-U.S.	f > m	global	.532

Note. Racial composition of United States samples were coded as composed of participants who were majority (75% or more) White, Black, Hispanic, Asian, racially diverse (no majority), or not reported (NR). Samples from countries other than the United States were coded as non-U.S. Year = year of publication; Pub = published or not; Age = mean age of the sample; *N*m = number of males; *N*f = number of females; f > m = more females than males; f = m = equal number of males and females; m > f = more males than females; lang = language; socsci = social sciences; *d* = biased effect size.
[a] Denotes an effect that would be considered an outlier based on the criteria proposed by Tabachnick and Fidell (2007).

courses), social sciences (including courses in humanities as well as social sciences), and others (including the few courses that could not be classified elsewhere). It is important to note that physical education marks were excluded as they were deemed to reflect physical rather than intellectual achievement. In addition, the global measure was included only when individual subject marks were not available. Therefore, when a study reported both separate subject marks and an overall GPA, only subject marks were included. Although mean age of the sample was noted as a sample-level variable, the categorical approach followed by Lindberg et al. (2010) was used here but with the source of the mark as a measure-level variable. Therefore, effect sizes were categorized as originating in elementary school, junior/middle school, high school, undergraduate, and graduate sources. Finally, the scale of measurement was coded as point scale, percent, letter, standardized, and others/not reported.

The coding of average age requires some clarifications as this value was not always reported by individual researchers. However, if the school grade was given, then the age variable was coded using the approach proposed by Voyer et al. (1995). For example, children in Grade 1 are typically 6 years old, whereas first-year undergraduate students are usually 19 years old. Following this approach, medical and graduate students were coded as having a mean age of 25. Finally, if an age range was reported, the midpoint was used.

To ensure the clarity and validity of coded variables, a coding sheet that included an entry for all coded variables was first prepared. Then, a subset of 80 studies (accounting for 145 effect sizes) was coded independently by the two authors, two experienced meta-analysts. This coding involved 15 variables (not all of them used in the moderator analysis), that is, sample ID (a crucial variable for multilevel analysis), year of publication, publication status (published or not), mean age of sample, type of school (public, private), sample school level (elementary, high school, etc.), racial composition, national origin, number of males, number of females, source of the grade, course content area, scale of measurement, statistic used, and effect size. Therefore, a total of 2,175 entries (15 variables × 145 effect sizes) produced only 18 disagreements, resulting in an interrater reliability of 99.2% (2,157 agreements/2,175 total entries; $\kappa = .983$). This high interrater reliability clearly reflects the relatively straightforward coding. Therefore, the second author coded the remaining entries in consultation with the first author. Specifically, in the few instances where data were unclear, coding proceeded by mutual agreement of the two authors.

Measure of effect size. The standardized mean difference in performance (Cohen's *d*; Cohen, 1988) was the effect-size measure (Borenstein et al., 2009). In the present context, the effect size was calculated as the mean for females minus that for males divided by the pooled standard deviation. This calculation reflects the assumption that women would outperform men based on the literature presented so far. Thus, positive effect sizes reflect a female advantage, and negative effect sizes reflect a male advantage. We calculated the effect sizes based on the formula presented by Cohen (1988) when means and standard deviations were available, which was the case for 286 out of the 502 effect sizes (57.0%). However, only an inferential statistic (typically *t* test, *p*, *r*, or *F*) was available in the remaining cases. In this situation, the formulae presented by Lipsey and Wilson (2001) were used. In fact, when a study reported letter grades, this was often the only available option. In all cases, the computations were completed using the effect-size calculator provided on David Wilson's webpage (http://mason.gmu.edu/~dwilsonb/downloads/ES_Calculator.xls). The small-sample correction proposed by Hedges and Becker (1986) was applied to all effect sizes, although the effect sizes reported in Table 1 do not reflect this correction. Finally, when an effect size was not significant but no specific mean or inferential statistics values were presented, we contacted authors of work published in English by e-mail with a request for more information. An e-mail address could not be found for 13 of 40 such authors. Among the remaining authors, eight replied but did not have access to the relevant data anymore, whereas an additional five authors provided usable information. The remaining effect sizes ($k = 48$) were coded as zero following the approach recommended by Rosenthal (1991) as a conservative measure.

Data Analysis

The main purposes of the present analysis were to determine whether overall significant gender differences in school marks exist and to identify moderator variables affecting the magnitude or direction of these gender differences. The first goal was met by calculating an overall effect size and testing its significance. The second goal required the estimation of whether a given moderator has a significant effect on the magnitude of gender differences.

Multilevel meta-analysis. The comparison of effect sizes for marks across different course areas (global, math, science, language, social science, other) is possibly one of the most crucial component of the present analysis. However, many of the retrieved studies reported effect sizes for two or more content areas from the same participants, resulting in a sample that violates the assumption that effect sizes should be independent from each other (Borenstein et al., 2009). Analysis was further complicated by the fact that most studies reported effect sizes for a variable number of these content areas, with some providing relevant data for only one area whereas others covered many but not all areas. This precluded conduct of a multivariate meta-analysis (Becker, 2000; Kalaian & Raudenbush, 1996). However, the multilevel modeling approach to meta-analysis can handle such an unequal number of correlated effect sizes per study, and it makes no assumptions of independence among effect sizes (Marsh, Bornmann, Mutz, Daniel, & O'Mara, 2009). In fact, multilevel modeling in general is designed to analyze data that arise from such a nested structure (Raudenbush & Bryk, 2002). Furthermore, Van den Noortgate and Onghena (2003) claimed that multilevel models overpower fixed- and random-effects models due to their lack of independence assumption. Accordingly, we adopted the multilevel approach for analysis of the overall sample of effect sizes and for consideration of course content as a moderator.

In the present context, multilevel analysis was computed by analyzing data as organized in two levels: effect sizes nested within samples. This data structure resulted in 502 effect sizes relevant to measures (Level 1) nested within 369 samples (Level 2).

Mixed-effects meta-analysis. After demonstrating that course content was a significant moderator, effect sizes for each content area were then examined separately to identify significant moderator variables among the variables previously mentioned (gender composition, source of grades, national origin, racial composition, type of school, scale, and year of publication). Mixed-effects meta-analysis was therefore applied separately to each type of course material as these groupings were expected to include only independent effect sizes.[3] These analyses followed a mixed-effects approach to the meta-analysis analogue to analysis of variance for categorical moderators and to metaregression for the continuous moderator (namely, grand mean-centered year of publication), as recommended by Borenstein et al. (2009). Analyses were conducted by means of the SPSS macros prepared by Wilson (2005). Following the approach recommended by Borenstein et al., the effect sizes were weighted by their inverse variance. Interaction terms were not considered in data analysis as we could not justify them on the basis of the retrieved literature. Furthermore, as many variable combinations did not exist in the data set (e.g., majority Hispanic samples in graduate school sources), a proper test of interactions, even for exploratory purposes, would not be possible.

Results

As a preliminary data analysis, outliers were first identified as effect sizes that were more than 3.29 standard deviations away from the grand mean as proposed by Tabachnick and Fidell (2007). Only two such outliers were identified. However, considering the large number of effect sizes retrieved here, a few outliers are to be expected (Tabachnick & Fidell, 2007). Furthermore, closer examination of the source of these outliers suggests that the relevant studies are not clearly problematic otherwise. Accordingly, data analyses were conducted on the full sample of retrieved effect sizes. The outliers are nevertheless denoted by a superscript (i.e., ^{a}) in Table 1 for readers who might want more details on their

---

[3] Upon closer examination of the individual samples of studies, it turned out that only the science courses grouping was formed exclusively of independent effect sizes. For language, math, social sciences, and other courses, the few nonindependent effect sizes were averaged, following one of the commonly recommended approaches (Borenstein et al., 2009). For global measures, nonindependence arose because of multiple sources of marks from the same sample for 13 samples. We kept these samples as is to allow a proper test of this moderator, although this nonindependence should be kept in mind when interpreting the results for other moderators in global measures. The averaging of effect sizes accounts for the discrepancies in sample sizes for each course when comparing Tables 2 and 3.

characteristics. The final sample was therefore composed of 502 effect sizes drawn from 369 independent samples. This final sample of studies combined results obtained with 538,710 males and 595,332 females.

Multilevel Meta-Analysis

The multilevel data analysis was used to compute the overall effect and to determine whether course content area moderated the magnitude of effect sizes significantly. Analyses were accomplished with HLM Version 7 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011). HLM 7 is a standard program for analyzing hierarchical or nested data. The estimation procedure relied on a full-maximum-likelihood model with the effect sizes treated as random effects and moderators treated as fixed effects. This software uses an approach to variance–covariance matrix computation akin to an unstructured matrix (Garson, 2013) so that no assumptions are made about the type of matrix involved. As course content is a categorical variable, it was dummy-coded. Using this method, the intercept represents the estimated effect size for the reference category, and its test of significance examines whether it is significantly different from zero. The coefficient for other categories reflects the difference between their effect size and the one observed in the reference category. The corresponding test of significance examines whether each grouping is significantly different from the reference category (Pedhazur, 1997). As is typical in multilevel meta-analysis (Raudenbush & Bryk, 2002), the Level 1 (measures) sampling variance was assumed to be known (as represented in the variance calculated for each effect size; see Borenstein et al., 2009), whereas Level 2 (sample) variance was estimated in each model. Accordingly, the multilevel model used in the present meta-analysis is considered a V-known model and results in precision weighted effect-size estimates (Raudenbush & Bryk, 2002).

Is there an overall significant gender difference? Results showed a significant female advantage on school marks, reflecting an overall estimated *d* of 0.225 (95% CI [0.201, 0.249]). As the confidence interval did not include zero, the overall effect size is significant with $p < .05$. This finding was established by considering the results of the null model (intercept only) and testing the significance of the observed effect size (Raudenbush & Bryk, 2002).

In view of the fact that we had to code 48 effect sizes (9.6% of effect sizes) as zero in the final sample due to missing information, one could argue that the overall results reported so far reflect a lower-bound estimate of the true effect. Accordingly, the overall estimated mean effect size was also computed on the basis of the known effects only, with this value viewed as an upper-bound estimate of the true effect. Therefore, this analysis included 454 effect sizes from 332 samples and produced a mean estimated effect size of 0.251 (95% CI [0.225, 0.277]), still indicating an overall female advantage in school marks. Interpreted strictly, the difference between the lower- and upper-bound estimates suggests that the true effect size might be approximately 0.238. Nevertheless, further analyses proceeded with the original sample (including effect sizes coded as zero) to provide a complete analysis of the retrieved research.

Is course content area a significant moderator? The examination of course content as a potential moderator relied on determining whether it accounted for significant variance in the retrieved effect sizes compared to the null model. This analysis showed that course contributed significantly to variance in effect sizes, $\chi^2(5) = 2{,}007.96$, $p < .001$. As seen in Table 2, all effect sizes were in the direction of a female advantage, and none of the 95% CIs included zero. Furthermore, the largest effects were observed for language courses, and the smallest gender differences were obtained in math courses. The analysis also showed that the effect sizes were significantly larger in language marks than in the reference category, global measures, $t(128) = 4.19$, $p < .001$. In contrast, the magnitude of gender differences was significantly smaller in mathematics, science, and social sciences courses when compared to global measures, smallest $t(128) = -2.45$, $p = .016$. Gender differences in *other* courses and global measures were statistically similar ($p > .51$).

Moderator Analyses

Having established differences in overall magnitude of gender differences among the different content areas, moderator analysis examined the potential influence of the remaining moderators separately for each course content area. Therefore, all moderators (gender composition, source of grades, national origin, racial composition, type of school, scale, and year of publication) were examined systematically to determine if they produced significant between-group heterogeneity (for categorical variables) or accounted for significant variance (for the only continuous moderator, year of publication grand mean-centered). Only moderators significant at the .05 level are presented here. Furthermore, multiple comparisons among effect sizes followed the *z*-score method outlined by Borenstein et al. (2009) at the .05 level of significance. Note that variability across degrees of freedom for a given moderator is due to the fact that not all categories are represented across content areas. Results of the moderator analysis are summarized in Table 3.

The moderator analysis showed that source of the marks was a significant moderator for global measures, $\chi^2(4) = 52.23$, $p < .001$; language, $\chi^2(4) = 17.71$, $p < .001$; math, $\chi^2(4) = 20.08$, $p < .001$; and science courses, $\chi^2(4) = 8.20$, $p = .042$. Multiple comparisons showed that, for global measures, effect sizes were significantly larger for all sources compared to graduate school sources. For global measures, language, and science courses, effect sizes were larger in junior/middle school and high school than in college. For language and math, gender differences were smaller in elementary school than in junior/middle and high school. Finally,

Table 2
Results for Course Material in the Multilevel Analysis

Moderator	Sample size (*k*)	Estimated mean *d*	95% confidence interval
Course material			
Global measures	258	0.249	0.217, 0.281
Language	81	0.374	0.316, 0.432
Mathematics	93	0.069	0.014, 0.124
Science	31	0.154	0.078, 0.230
Social sciences	26	0.174	0.117, 0.231
Other	13	0.285	0.175, 0.395

for math, effect sizes were smaller in elementary school than in college.

In addition, national origin was a significant moderator for global measures, $\chi^2(2) = 28.63$, $p < .001$; language, $\chi^2(2) = 23.19$, $p < .001$; and math courses, $\chi^2(2) = 23.56$, $p < .001$. Here, multiple comparisons showed that Scandinavian samples produced smaller gender differences than North American samples and the rest of the world in global measures. Language courses showed smaller effects for the rest of the world compared to North American and Scandinavian samples. Finally, math courses produced smaller effect sizes for the rest of the world compared to North American samples

Racial composition was also a significant moderator for global measures, $\chi^2(5) = 15.04$, $p = .010$; language, $\chi^2(4) = 12.24$, $p = .016$; and math courses, $\chi^2(4) = 20.78$, $p < .001$. In all cases, the non-U.S. category produced the smallest magnitude of gender difference. For math courses, the non-U.S. category was significantly smaller than for all other racial composition categories. For language courses, the only significant difference was between non-U.S. samples and those where the racial composition was not reported. Finally, for global measures, non-U.S. samples produced smaller gender differences than all other categories except samples with a majority of Hispanic participants and those with a diverse racial composition. It should be noted that when non-U.S. samples were removed from the data, racial composition did not emerge as a significant moderator of effect sizes in any of the course content areas (all *p*s > .4).

Finally, gender composition had a significant influence on the magnitude of gender differences in math, $\chi^2(2) = 12.09$, $p = .002$, and science courses, $\chi^2(3) = 8.19$, $p = .042$. In both cases, the effect size did not significantly differ from zero in samples with more males than females, although for math, the difference was significantly smaller only when compared with samples with more females than males. In science courses, all differences among effect sizes were significant, with the following order: (equal number of females and males) > (more females than males) > (more males than females).

No other moderators achieved significance with $p < .05$ (all *p*s > .11). Therefore, type of school (private, public), grading scale (point scale, percent, letter, standardized, others/not reported), and year of publication did not emerge as significant moderators in all analyses. However, considering the importance of year of publication for claims relevant to the boy crisis and to circumvent the argument that insufficient power was achieved to detect the effect within each course content area, this variable was examined (grand mean-centered) in the full sample by means of multilevel modeling. This analysis also revealed nonsignificant findings ($p > .16$) and a very small negative coefficient of −0.001. It is therefore clear that, in the present sample, the magnitude of gender differences in school marks has remained stable in the years covered here (from 1914 to 2011). Finally, no significant moderators were identified for social sciences and *other* courses.

Publication Bias

We examined the possibility of a publication bias by comparing the mean estimated effect sizes for the published ($k = 339$ samples) and unpublished research ($k = 30$ samples). This analysis showed a significant influence of publication status, $\chi^2(1) = 5.97$, $p = .014$. This finding showed that effect sizes derived from

Table 3

Results of the Moderator Analysis as a Function of Course Content Area

	Global		Language		Math		Science	
Moderator	*k*	*d* (95% CI)	*k*	*d* (95% CI)	*k*	*d* (95% CI)	*k*	*d* (95% CI)
Source of marks								
Elementary	24	0.37 (0.27, 0.47)	23	0.20 (0.10, 0.30)	28	−0.01 (−0.08, 0.07)	9	0.10 (−0.02, 0.22)
Junior/middle	23	0.36 (0.26, 0.45)	20	0.45 (0.35, 0.56)	25	0.23 (0.15, 0.30)	9	0.23 (0.12, 0.34)
High school	51	0.39 (0.33, 0.46)	17	0.47 (0.35, 0.59)	17	0.11 (0.01, 0.20)	5	0.16 (0.01, 0.31)
College/university	131	0.21 (0.17, 0.25)	7	0.21 (0.01, 0.40)	20	0.12 (0.04, 0.20)	8	0.01 (−0.11, 0.12)
Graduate school	28	0.06 (−0.02, 0.15)	—	—	3	0.25 (−0.01, 0.50)	—	—
National origin								
North America	198	0.29 (0.25, 0.32)	37	0.47 (0.39, 0.56)	63	0.17 (0.13, 0.22)		
Scandinavia	22	0.03 (−0.06, 0.12)	8	0.35 (0.18, 0.51)	7	0.03 (−0.10, 0.15)		
Rest of the world	37	0.27 (0.19, 0.34)	22	0.15 (0.04, 0.25)	22	−0.04 (−0.11, 0.04)		
Racial composition								
Majority White	41	0.29 (0.21, 0.36)	6	0.37 (0.17, 0.59)	4	0.13 (0.04, 0.22)		
Majority Black	20	0.36 (0.25, 0.47)	2	0.50 (0.13, 0.87)	3	0.35 (0.13, 0.57)		
Majority Hispanic	3	0.32 (0.25, 0.47)	—	—	—	—		
Racially diverse	43	0.24 (0.17, 0.32)	5	0.39 (0.17, 0.62)	9	0.17 (0.04, 0.30)		
Other/NR	85	0.29 (−0.03, 0.24)	20	0.50 (0.38, 0.62)	31	0.18 (0.12, 0.24)		
Non-U.S.	65	0.17 (0.11, 0.23)	34	0.25 (0.16, 0.33)	35	0.01 (−0.05, 0.07)		
Gender composition								
F > M					48	0.16 (0.11, 0.21)	19	0.14 (0.07, 0.22)
F = M					9	0.15 (0.03, 0.26)	3	0.32 (0.17, 0.47)
M > F					35	0.03 (−0.02, 0.09)	9	0.01 (−0.09, 0.10)

Note. The table presents sample size (*k*) and the mean weighted *d* for each moderator category with the 95% confidence interval (CI) in parentheses. The mean weighted effect size is significantly different from zero with $p < .05$ if the 95% CI for *d* does not include zero. Dashes indicate that a category is not represented in the data. Junior/middle = junior/middle school; NR = not reported; F > M = more females than males; F = M = equal number of males and females; M > F = more males than females.

Table 4

Summary of Results for the Significant Moderators Observed in the Present Meta-Analysis

Moderator	Pattern of results
Overall	Overall *d* of 0.225 found to be significantly larger than zero, reflecting an overall female advantage.
Course material	Effect sizes were larger in language marks than in global measures but smaller in mathematics, science, and social sciences courses than for global measures.
Source of the marks	Significant in global measures, language, math, and science courses. The female advantage was not significant for graduate school samples on global measures and in elementary school for math. Largest effects are typically found in junior/middle and high school.
National origin	Significant in global measures, language, and math courses. Generally, Scandinavian samples produced smaller gender differences than North American samples but only for global measures and math.
Racial composition	Significant in global measures, language, and math courses. Not significant if non-U.S. samples are removed. Generally, a majority of Black students produced large effect sizes. However, the findings are driven by smaller gender differences in non-U.S. samples.
Gender composition	Significant in math and science courses. The gender difference was not significant in samples where there were more males than females.

published studies (estimated mean $d = 0.216$, 95% CI [0.191, 0.241]) were significantly smaller than those obtained in unpublished studies (estimated mean $d = 0.328$, 95% CI [0.246, 0.410]).[4]

Discussion

The purpose of the present analysis was to provide a summary of findings pertaining to gender differences in scholastic achievement as measured by teacher-assigned school marks. Quantification of these gender differences as well as the identification of relevant moderator variables formed the primary goals of the analysis. From an exhaustive search and our examination of the literature, we felt that such a comprehensive analysis was missing and that it would complement meta-analyses that focused exclusively on achievement tests results.

The present analysis relied on two complementary approaches to meta-analysis, with a multilevel model for the whole sample and a mixed-effects model analysis within content areas. The findings are summarized in Table 4. Potentially the most crucial finding of the present analysis is that the female advantage in the overall sample was significantly larger than zero. In addition, the results also showed that the largest gender differences were found in language courses and the smallest were in math courses, although all were significantly larger than zero.

The moderator analysis, conducted separately for each course content area, revealed four significant moderators although their effects were confined to courses in language, math, science, and global measures. Specifically, source of the marks was significant in all four of these areas, while national origin and racial composition were significant for global measures, language, and math courses. Finally, gender composition was significant for math and science courses. It is noteworthy that year of publication was not a significant moderator in any of the analyses. In addition, no significant moderators could be identified for social sciences and for *other* courses.

Overall Results: Implications

The most important finding observed here is that our analysis of 502 effect sizes drawn from 369 samples revealed a consistent female advantage in school marks for all course content areas. In contrast, meta-analyses of performance on standardized tests have reported gender differences in favor of males in mathematics (e.g., Else-Quest et al., 2010; Hyde et al., 1990; but see Lindberg et al., 2010) and science achievement (Hedges & Nowell, 1995), whereas they have shown a female advantage in reading comprehension (e.g., Hedges & Nowell, 1995). This contrast in findings makes it clear that the generalized nature of the female advantage in school marks contradicts the popular stereotypes that females excel in language whereas males excel in math and science (e.g., Halpern, Straight, & Stephenson, 2011). Yet the fact that females generally perform better than their male counterparts throughout what is essentially mandatory schooling in most countries seems to be a well-kept secret considering how little attention it has received as a global phenomenon. In fact, the popular press and elected representatives still focus on findings that fit stereotypical expectations (e.g., Tyre, 2006), despite an attempt over 20 years ago by Kimball (1989) to bring the female advantage to light for mathematics. The present findings bolster Kimball's results and extend them to other content areas in school.

Several plausible explanations align with the overall effect in school marks. First, a number of sociocultural factors, such as expectancy and value (see Eccles et al., 1983), have been proposed. The expectancy-value model suggests that achievement behavior can be predicted by the expectancy for future success and value given to a task. Therefore, from the perspective of the expectancy-value model, if one has low expectancy of success and sees little future value in a specific course topic, one is not likely to be motivated to work hard in that course (Steinmayr & Spinath, 2008). From this perspective, expectancy-value could affect both how much students invest in school work and how much time and effort teachers invest in specific students. The general idea therefore would be that females do more poorly than males in math because it has low expectancy and value for them, whereas this would be reversed for language courses. However, this model cannot fully explain our finding that females have an advantage in

---

[4] In addition, the Egger, Davey Smith, Schneider, and Minder (1997) test provided a further examination of publication bias. Results of this analysis supported the presence of a publication bias only for global measures (intercept = 2.66, 90% CI [1.85, 3.47]) and math courses (intercept = 1.05, 90% CI [0.49, 1.61]). The trim-and-fill correction, computed with MIX 2.0 Pro (Bax, 2011) for these two categories, suggested the presence of a strong bias for math courses as the adjusted effect size was near zero (corrected mean $d = 0.025$, 95% CI [0.015, 0.038]). However, for global measures, the procedure did not deem trim and fill necessary. This suggests that the finding of an apparent publication bias for global measures might reflect the operation of factors other than publication bias (Egger et al., 1997).

all course material. Therefore, more refined suggestions drawing from the expectancy-value model are considered in the context of other accounts of the female advantage in school marks.

Social factors, such as the fact that parents tend to attribute math performance to abilities for males and to efforts for females (Eccles, Jacobs, & Harold, 1990) might also lead males and females to approach school work differently (Kenney-Benson et al., 2006). Specifically, the differential attributions made by parents might lead them to encourage more effort in females than in males, at least in math courses. This differential amount of encouragement could account in part for the slight female advantage in math. In fact, findings that parents encourage efforts on school work more in females than in males for all content areas would account for the generalized female advantage in school marks. Research by Varner and Mandara (2013) confirmed this possibility in African American adolescents. In fact, their findings led them to conclude that "reducing gaps in parenting may help reduce the gender gap in achievement" (Varner & Mandara, 2013, p. 12). In contrast, Raymond and Benbow (1989) reported that differential encouragement is guided more by the perceived talent domain of a child than by gender. In view of these contradictory findings, the possibility of gender-differentiated parental encouragement toward school achievement might be an important avenue for more research.

The possible influence of stereotype threat is another social factor that has been suggested recently as a possible explanation of gender difference in school achievement. Stereotype threat occurs when a group's performance is affected by the knowledge that its members belong to a social group that is not expected to perform well in a task. In this context, Hartley and Sutton (2013) found that even at an early age, both girls (starting at age 4) and boys (starting at age 7) hold the belief that adults expect girls to be better students than boys. Hartley and Sutton then proceeded to show that emphasizing or countering this belief had a negative or positive effect, respectively, on boys' reading, writing, and math performance. These manipulations did not affect girls' performance. The study conducted by Hartley and Sutton supports the possibility that stereotype threat might affect expectancy for success, which in turn might affect effort and persistence in the classroom. These speculations suggest that more research on the influence of stereotype threat in the classroom could be fruitful.

Gender differences in learning styles (Dweck, 1986) could also be seen as having broad applicability as they would be relevant to all course areas. According to Kenney-Benson et al. (2006), the learning style of females tends to emphasize mastery over performance in task completion, whereas males tend to show the reverse emphasis. Mastery emphasis means that one pursues work in the hope of understanding the material, whereas performance emphasis indicates a focus on one's marks. When the findings of gender differences in mastery/performance emphasis are considered in the context that mastery emphasis generally produces better marks than performance emphasis, this could account in part for males' lower marks than females.

However, reports of gender differences in mastery/emphasis are contradictory (see Patrick, Ryan, & Pintrich, 1999). Therefore, this factor might not provide a solid account of the generalized female advantage in school marks.

Biological influences have also been proposed as possible factors of relevance, as they could underlie gender differences in activity levels (generally higher in males; Campbell & Eaton, 1999). This factor would potentially make it easier for females than for males to pay attention in class (Kenney-Benson et al., 2006). Gender differences in activity level might also relate to temperamental gender differences (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). Specifically, the meta-analysis conducted by Else-Quest et al. (2006) suggested a female advantage in effortful control and a male advantage in surgency. Taken together, these gender differences in activity levels and temperament could manifest themselves in the classroom. For example, gender differences in class behavior could affect teachers' subjective perceptions of students, which in turn might affect their grades (Bennett, Gottesman, Rock, & Cerullo, 1993). This subjective component of school marks should not be overlooked as it has been shown to affect teachers' evaluation of their students, potentially leading to sex-biased treatment (Chalabaev, Sarrazin, Trouilloud, & Jussim, 2009) and self-fulfilling prophecies (Jussim, Robustelli, & Cain, 2009).

This nonexhaustive list of explanations for the female advantage in school marks emphasizes the complexity of this issue. In reality, a multitude of factors might account for the female advantage in school marks. Some of these explanations are revisited in the context of the moderator analysis results.

Practical significance. Implications of the overall magnitude of the observed effects require some discussion. Admittedly, the gender difference observed here would be classified as small based on Cohen's (1988) categorization of effect sizes (i.e., values of 0.2 and lower are considered a small effect size). In fact, when considering the magnitude of the effects, it might be tempting to conclude that many of the estimates presented in Tables 2 and 3 are so small as to reflect nonexistent gender differences. However, the findings are striking in their consistency. Specifically, only two of the mean effect sizes presented in the results tables reflect a negative effect (both in Table 3, for math in elementary school samples and in *rest of the world*). Even then, in both cases, the confidence intervals include zero. Therefore, none of the results indicate a male advantage. The consistency of the results can also be considered in the context of Abelson's (1985) argument that apparently small effects should not be overlooked as the weight of accumulated evidence also has to be considered, especially when this evidence has a cumulative effect. Essentially, the female advantage on school marks is relatively small, but it is a common finding in the literature. Closer to Abelson's argument, as males start obtaining lower grades than females early in schooling, this might have a cumulative effect in the long run with school marks (regardless of their magnitude) potentially affecting future behavior each step along the way. In fact, Abelson mentioned educational interventions as one example of potentially cumulative processes.

To put the present findings in perspective, an effect size of 0.225 would reflect approximately a 16% nonoverlap between distributions of males and females (Cohen, 1988). Thus, a crude way to interpret this finding is to say that, in a class of 50 female and 50 male students, there could be eight males who are forming the lower tail of the class marks distribution. These males would be likely to slow down the class, for example, and this could have cumulative effects on their school marks. Of course, this is not a completely accurate way to interpret the nonoverlap, but it should serve to illustrate the importance of this finding. By comparison, considering values obtained in achievement tests, the overall *d* of 0.11 (in favor of females) reported by Hyde and Linn (1988) for verbal tests would reflect a nonoverlap of about 8.5%, whereas the overall *d* of 0.05 (in favor of males) reported by Lindberg et al. (2010) for math tests indicates a nonoverlap of about 4%. The present findings should therefore not be qualified as representing a trivial effect.

With this in mind, the moderator analysis has uncovered a number of factors that affect the magnitude of gender differences. As such, these factors may serve as a guide for further research. Accordingly, we now focus on the information that can be derived from the significant moderators we identified.

Moderator Analysis: Implications

Course material as moderator. Essentially, the examination of course material as a moderating factor showed that the largest effects were obtained in language courses, whereas the smallest effects were in math and science courses. This should not come as a surprise based on achievement tests results, although the direction of the effect for marks (always female advantage in Table 2) contradicts tests results and stereotypical expectations. When accounting for the variability in magnitude of effect sizes across content areas, it is possible that gender differences in interests (Su, Rounds, & Armstrong, 2009) could lead males and females to different expectancy-value on various course content areas, which in turn might produce gender-related fluctuations on the level of motivation in these different course areas (Steinmayr & Spinath, 2008). Considered together, these factors could explain in part why the female advantage is smallest in math and science (see Table 2), areas that are stereotypically viewed as masculine (Nosek et al., 2009).

Source of the marks as moderator. Findings relevant to source of the marks showed that this variable had a significant influence for global measures, as well as language, math, and science courses. Interestingly, the pattern of results suggested fairly stable effect sizes across schooling, except at the graduate level for global measures. In language, math, and science courses, effect sizes decreased between high school and university. In contrast, the magnitude of gender differences increased from elementary to middle school in these three course areas.

The finding for graduate school students can be interpreted in a straightforward manner as it would essentially reflect samples of highly motivated, top-quality students who are working in (presumably) their favorite field. Therefore, most expectancy-value and motivational factors would be equated, and this would potentially leave little room for fluctuation in grades. This limited amount of fluctuation would likely restrict the range of marks obtained. This in turn would reduce the likelihood of obtaining significant gender differences. Therefore, this statistical factor should not be overlooked. In interpreting gender differences in school marks for graduate courses, it should be noted that, although a relatively large effect was observed in graduate school for math, it did not significantly differ from zero (as seen in the 95% CI; see Table 3).

The results in math and science suggest that the female advantage in school performance in these areas does not emerge until junior/middle school. Unfortunately, the retrieved data do not provide information that might help elucidate the reasons for this relatively late emergence of the female advantage in these domains. Empirical work is therefore required to determine which factors contribute to the absence of sex differences in math and science at the elementary school level.

National origin as moderator. National origin was considered as three categories (North America, Scandinavia, other countries) based on a post hoc consideration of nationalities represented in the final sample. The fact that it is a significant moderator of gender differences in global measures as well as in language and math courses constitutes a somewhat serendipitous finding of the present analysis. Specifically, we found that Scandinavian countries had significantly smaller effects than North American samples in global measures and math courses, with language courses following this trend. The placement of the *other countries* category varied across areas. In fact, *other countries* even produced a nonsignificant negative effect for math. In hindsight, such findings relevant to national origin should not be surprising considering that Scandinavian countries rate high on the measures of gender equity examined by Else-Quest et al. (2010). As a matter of fact, Else-Quest et al. also reported that gender differences in math achievement tests are reduced in societies where there is more gender equity. However, it is rather difficult to apply directly the explanations offered by Else-Quest et al. as they emphasized societal factors that might account for poor female performance in math achievement tests, whereas we are concerned here with males' poorer performance compared to females. It is also important to consider that, despite the scope of the present study, a limited sample of research in Scandinavian countries was included (see Table 3). In addition, gender differences in language courses are still sizable in Scandinavian samples. Therefore, gender differences in school performance have not been eliminated even in these putatively gender equitable countries.

Racial composition as moderator. Racial composition was examined based on the notion that gender differences in achievement are larger when racial origin is considered as opposed to gender in United States samples (Mead, 2006). The working hypothesis for this moderator was that gender differences would be largest among samples composed of a majority of Black students. However, a crucial finding for this moderator was that, when only United States samples were considered in data analysis, racial composition did not contribute significantly to variance in the magnitude of gender differences. Nevertheless, racial composition did prove to be a significant moderator for global measures as well as for language and math courses when non-U.S. samples were included in the data. In addition, in all cases, samples with a majority of Black students produced the largest effect size (see Table 3), but this finding reflected only a nonsignificant trend. The only significant differences among effect sizes were typically due to the smaller gender differences observed in non-U.S. samples. This finding suggests that gender differences in school marks are particularly pronounced and relatively homogeneous in United States samples. However, the present data do not allow speculations on factors that might account for this homogeneity.

Gender composition as moderator. Before considering the influence of gender composition as a moderator of gender differences in school marks, it is important to remember that it was used as an indirect way to determine whether a field of study was female or male dominated. With this in mind, this moderator achieved significance in math and science only, and the results showed that samples with more males than females produced no significant gender differences in school marks. This finding suggests that the female advantage is reduced in male-dominated fields. Although this pattern of results provides a manipulation check on how we conceptualized gender composition, it does not shed much light on their cause. Empirical research would be required to determine the specific factors that might account for the present findings relevant to gender composition. It should also be noted that the gender difference was not necessarily larger in samples where there were more females than males. This serves as a reminder that the gender composition variable as operationalized here is only a crude measure of the underlying concept.

Year of publication as moderator. Although year of publication was not a significant moderator, it requires some discussion in the context of a potential boy crisis. Specifically, boy-crisis proponents suggest that males have started lagging behind females in terms of school achievement only recently (Tyre, 2006). In our analysis, support for the claim of a boy crisis required findings of a significant positive relation between year of publication and magnitude of gender differences. This claim was not supported by the results for each content area and for the whole sample. Therefore, the data in the present sample, ranging in years from 1914 to 2011, suggest that boys have been lagging for a long time and that this is a fairly stable phenomenon. Accordingly, it might be more appropriate to claim that the boy crisis has been a long-standing issue rather than a recent phenomenon.

Other courses and social sciences courses. Finally, no significant moderators emerged for the other courses and social sciences courses categories. In reality, a potential reason for this finding is that these categories consisted of highly heterogeneous course content. Therefore, the large amount of extraneous variance involved as a result of this heterogeneity might have precluded the emergence of a systematic influence for the moderators considered here. Nevertheless, these two groupings also produced a significant overall female advantage for school marks, testifying to the robustness of this advantage.

Limitations

The present meta-analysis answered a number of questions pertaining to the magnitude of gender differences in school achievement and some of the factors that moderate them. However, as in all undertakings of this magnitude, some limitations should be discussed.

At the methodological level, it was not always possible to include or operationalize factors in a satisfactory way. The most striking example is the operationalization of SES as whether a sample originated in public or private school. It is not particularly surprising that it did not produce significant findings considering that, for example, even students from a family with a high SES could attend public school. The reverse is also possible, for example, if a student from a low SES family received an entrance scholarship to a private school. The unavailability of this information in many of the studies was an obstacle to a better operationalization of this variable. Therefore, no strong claims can be derived from the present results concerning the potential influence of SES on gender differences in school marks.

A further methodological aspect with statistical ramifications concerns the decision to code nonsignificant effect sizes as having a value of zero when no data could be obtained from authors. Computation of the overall effect sizes while excluding the effect sizes coded as zero in this manner produced a slightly larger upper-bound estimate of 0.251 (as opposed to a lower-bound value of 0.225 when such effect sizes were included). However, as the confidence interval for this upper-bound estimate overlapped with the original value, they can be seen as falling in the same range of estimates. Therefore, inclusion of effect sizes coded as zero in all analyses likely contributed to obtaining a better reflection of the true effect size in the population as it allowed consideration of all available studies.

The examination of a sample composed mostly of published research in the present analysis could be seen as biasing the results. However, the finding that the unpublished research included here typically reported larger gender differences than those we found in published work raises the possibility that reliance on published work might underestimate the magnitude of gender differences in school marks. Accordingly, any evidence of a publication bias in the present sample could be due to the fact that gender differences in favor of females are simply more prevalent than those in favor of males when considering school marks.

Conclusions

The present meta-analysis used two complementary analytic approaches to address questions pertaining to the existence of gender differences in school achievement as measured by teacher-assigned marks and the factors that moderate them. The results showed that these gender differences favored females in all fields of study. The effect sizes were generally small in magnitude, but their consistency suggests that they should not be ignored. The finding that the female advantage in school marks has remained stable across the years in the data retrieved (from 1914 to 2011) deserves emphasis as it contradicts claims of a recent boy crisis in school achievement.

The present article has laid the groundwork to establish the existence of a generalized female advantage in school achievement. However, much research is still required to determine factors related to gender differences in school performance as well as their possible causes.

References

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

*Abbott, A. A. (1981). Factors related to third grade achievement: Self-perception, classroom composition, sex and race. *Contemporary Educational Psychology, 6,* 167–179. doi:10.1016/0361-476X(81)90046-1

*Abedi, J. (1991). Predicting graduate academic success from undergraduate academic performance: A canonical correlation study. *Educational and Psychological Measurement, 51,* 151–160. doi:10.1177/0013164491511014

Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin, 97,* 129–133. doi:10.1037/0033-2909.97.1.129

*Adams, D., Astone, B., Nunez-Wormack, E., & Smodlaka, I. (1994). Predicting the academic achievement of Puerto Rican and Mexican-American ninth-grade students. *Urban Review, 26,* 1–14. doi:10.1007/BF02354854

*Adams, R. E., & Laursen, B. (2007). The correlates of conflict: Disagreement is not necessarily detrimental. *Journal of Family Psychology, 21,* 445–458. doi:10.1037/0893-3200.21.3.445

*Ajiboye, J. O., & Tella, A. (2006). Class attendance and gender effects on undergraduate students' achievement in a social studies course in Botswana. *Essays in Education, 18,* 1–11.

*Alessandri, S. M. (1992). Effects of maternal work status in single-parent families on children's perception of self and family and school achievement. *Journal of Experimental Child Psychology, 54,* 417–433. doi:10.1016/0022-0965(92)90028-5

*Alfan, E., & Othman, M. N. (2005). Undergraduate students' performance: The case of University of Malaya. *Quality Assurance in Education, 13,* 329–343. doi:10.1108/09684880510626593

*Al-Mannie, M. (1990). A study of selected factors and their relationship to academic performance of students at King Saud University. *College Student Journal, 24,* 173–183.

*Alnabhan, M., Al-Zegoul, E., & Harwell, M. (2001). Factors related to achievement levels of education students at Mu'tah University. *Assessment & Evaluation in Higher Education, 26,* 593–604. doi:10.1080/02602930120093913

*Altschul, I., Oyserman, D., & Bybee, D. (2006). Racial-ethnic identity in mid-adolescence: Content and change as predictors of academic achievement. *Child Development, 77,* 1155–1169. doi:10.1111/j.1467-8624.2006.00926.x

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: Macmillan.

*Anderson, J. L. (2006). Predicting final GPA of graduate school students: Comparing artificial neural networking and simultaneous multiple regression. *College and University, 81*(4), 19–29.

*Annor, P. (2010). Factors that affect the academic success of foreign students at Cardinal Stitch University. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 71*(6), 1925.

*Ari, A., Atalay, O. T., & Aljamhan, E. (2010). From admission to graduation: The impact of gender on student academic success in respiratory therapy education. *Journal of Allied Health, 39,* 175–178.

*Arrison, E. G. (1998). Academic self-confidence as a predictor of first year college student quality of effort and achievement. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 59*(3), 0742.

*Attaway, N. M., & Bry, B. H. (2004). Parenting style and Black adolescents' academic achievement. *Journal of Black Psychology, 30,* 229–247. doi:10.1177/0095798403260720

*Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences, 20,* 427–435. doi:10.1016/jlindif.2010.06.003

*Ayers, J. B., Bustamante, F. A., & Campana, P. J. (1973). Prediction of success in college foreign language courses. *Educational and Psychological Measurement, 33,* 939–942. doi:10.1177/001316447303300426

*Ayers, J. B., & Quattlebaum, R. F. (1992). TOEFL performance and success in a masters program in engineering. *Educational and Psychological Measurement, 52,* 973–975. doi:10.1177/0013164492052004021

*Azen, R., Bronner, S., & Gafni, N. (2002). Examination of gender bias in university admissions. *Applied Measurement in Education, 15,* 75–94. doi:10.1207/S15324818AME1501_05

*Babaoye, M. S. (2000). Input and environmental characteristics in student success: First-term GPA and predicting retention at an historically Black university. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 61*(12), 4669.

*Balsa, A. I., Giuliano, L. M., & French, M. T. (2011). The effects of alcohol use on academic achievement in high school. *Economics of Education Review, 30,* 1–15. doi:10.1016/j.econedurev.2010.06.015

*Banducci, R. (1967). The effect of mother's employment on the achievement, aspirations, and expectations of the child. *Personnel & Guidance Journal, 46,* 263–267. doi:10.1002/j.2164-4918.1967.tb03182.x

*Banreti-Fuchs, K. M. (1972). Attitudinal and situational correlates of academic achievement in young adolescents. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 4,* 156–164. doi:10.1037/h0082299

*Baucal, A., Pavlovic-Babic, D., & Willms, J. D. (2006). Differential selection into secondary schools in Serbia. *Prospects, 36,* 539–546. doi:10.1007/s11125-006-9011-9

Bax, L. (2011). MIX 2.0: Professional software for meta-analysis in Excel (Version 2.0.1.4) [Computer software]. Retrieved from http://www.meta-analysis-made-easy.com

*Bean, J. P., & Bradley, R. K. (1986). Untangling of satisfaction-performance relationship for college students. *Journal of Higher Education, 57,* 393–412. doi:10.2307/1980994

*Beaudin, B. Q., Horvath, J., & Wright, S. P. (1992). Predicting freshman persistence in economics: A gender comparison. *Journal of the Freshman Year Experience, 4,* 69–84.

Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499–525). doi:10.1016/B978-012691360-6/50018-5

*Beecher, M., & Fischer, L. (1999). High school courses and scores as predictors of college success. *Journal of College Admission, 163,* 4–9.

*Beer, J. (1989). Relationship of divorce to self-concept, self-esteem, and grade point average of fifth and sixth grade school children. *Psychological Reports, 65,* 1379–1383. doi:10.2466/pr0.1989.65.3f.1379

*Behrens, L. G., & Vernon, P. E. (1978). Personality correlates of over-achievement and under-achievement. *British Journal of Educational Psychology, 48,* 290–297. doi:10.1111/j.2044-8279.1978.tb03015.x

*Benedict, M. E., & Hoag, J. (2004). Seating location in large lectures: Are seating preferences or location related to course performance? *Journal of Economic Education, 35,* 215–231. doi:10.3200/JECE.35.3.215-231

*Bennett, R. E., Gottesman, R. L., Rock, D. A., & Cerullo, F. (1993). Influence of behavior perceptions and gender on teachers' judgments of students' academic skill. *Journal of Educational Psychology, 85,* 347–356. doi:10.1037/0022-0663.85.2.347

*Berndt, T. J., & Miller, K. E. (1990). Expectancies, values, and achievement in junior high school. *Journal of Educational Psychology, 82,* 319–326. doi:10.1037/0022-0663.82.2.319

*Betts, J. R., & Morell, D. (1999). The determinants of undergraduate grade point average: The relative importance of family background, high school resources, and peer group effects. *Journal of Human Resources, 34,* 268–293. doi:10.2307/146346

*Beyer, H. N. (1971). Effect of students' knowledge of their predicted grade point averages on academic achievement. *Journal of Counseling Psychology, 18,* 603–605. doi:10.1037/h0031757

Beyer, S. (1999). The accuracy of academic gender stereotypes. *Sex Roles, 40,* 787–813. doi:10.1023/A:1018864803330

*Bienstock, J. L., Martin, S., Tzou, W., & Fox, H. E. (2002). Medical students' gender is a predictor of success in the obstetrics and gynecology basic clerkship. *Teaching and Learning in Medicine, 14,* 240–243. doi:10.1207/S15328015TLM1404_7

*Bink, M. L., Biner, P. M., Huffman, M. L., Geer, B. L., & Dean, R. S. (1995). Attitudinal, college/course-related, and demographic predictors of performance in televised continuing education courses. *Journal of Continuing Higher Education, 43,* 14–20. doi:10.1080/07377366.1995.10400927

*Blackman, I., Hall, M., & Darmawan, I. G. (2007). Undergraduate nurse variables that predict academic achievement and clinical competence in nursing. *International Education Journal, 8,* 222–236.

*Blair, B. F., & Millea, M. (2004). Student academic performance and compensation: The impact of cooperative education. *College Student Journal, 38,* 643.

*Blaser, F. (1981). Factors related to success in a field experience general methods course. *Teacher Educator, 17*(3), 24–29. doi:10.1080/08878738109554790

*Boldt, K. E. (2000). Predicting academic performance of high-risk college students using Scholastic Aptitude Test scores and noncognitive variables. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 61*(6), 2181.

*Booth, R. (1983). An examination of college GPA, composite ACT scores, IQs, and gender in relation to loneliness of college students. *Psychological Reports, 53,* 347–352. doi:10.2466/pr0.1983.53.2.347

*Borde, S. F. (1998). Predictors of student academic performance in the introductory marketing course. *Journal of Education for Business, 73,* 302–306. doi:10.1080/08832329809601649

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis.* doi:10.1002/9780470743386

*Borup, J. H. (1971). The validity of American College Test for discerning potential academic achievement levels: Ethnic and sex groups. *Journal of Educational Research, 65,* 3–6.

*Bourquin, S. D. (1999). The relationship among math anxiety, math self-efficacy, gender, and math achievement among college students at an open admissions commuter institution. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 60*(3), 0679.

*Bowman, R. L., & Partin, K. E. (1993). The relationship between living in residence halls and academic achievement. *College Student Affairs Journal, 13,* 71–78.

*Brady, B. A., Tucker, C. M., Harris, Y. R., & Tribble, I. (1992). Association of academic achievement with behavior among Black students and White students. *Journal of Educational Research, 86,* 43–51. doi:10.1080/00220671.1992.9941826

*Bridgeman, B., & Lewis, C. (1994). The relationship of essay and multiple-choice scores with grades in college courses. *Journal of Educational Measurement, 31,* 37–50. doi:10.1111/j.1745-3984.1994.tb00433.x

*Bridgeman, B., & Wendler, C. (1991). Gender differences in predictors of college mathematics performance and in college mathematics course grades. *Journal of Educational Psychology, 83,* 275–284. doi:10.1037/0022-0663.83.2.275

*Britner, S. L. (2008). Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes. *Journal of Research in Science Teaching, 45,* 955–970. doi:10.1002/tea.20249

*Brog, M. J. (1985). Hemisphericity, locus of control, and grade point average among middle and high school boys and girls. *Perceptual and Motor Skills, 60,* 39–45. doi:10.2466/pms.1985.60.1.39

*Brooks, C. I. (1987). Superiority of women in statistics achievement. *Teaching of Psychology, 14,* 45. doi:10.1207/s15328023top1401_13

*Brooks, C. I., & Mercincavage, J. E. (1991). Grades for men and women in college courses taught by women. *Teaching of Psychology, 18,* 47–48. doi:10.1207/s15328023top1801_17

*Brooks, C. I., & Rebeta, J. L. (1991). College classroom ecology: The relation of sex of student to classroom performance and seating preference. *Environment and Behavior, 23,* 305–313. doi:10.1177/0013916591233003

*Brown, W. T., & Jones, J. M. (2004). The substance of things hoped for: A study of the future orientation, minority status perceptions, academic engagement, and academic performance of Black high school students. *Journal of Black Psychology, 30,* 248–273. doi:10.1177/0095798403260727

*Brubeck, D., & Beer, J. (1992). Depression, self-esteem, suicide ideation, death anxiety, and GPA in high school students of divorced and nondivorced parents. *Psychological Reports, 71,* 755–763. doi:10.2466/pr0.1992.71.3.755

*Buck, J. L. (1985). A failure to find gender differences in statistics achievement. *Teaching of Psychology, 12,* 100. doi:10.1207/s15328023top1202_13

*Bunce, J., & Calvert, B. (1974). Pupils' primary school record forms. *New Zealand Journal of Educational Studies, 9,* 42–51.

*Burgert, R. H. (1935). The relation of school marks to intelligence in secondary schools. *Journal of Applied Psychology, 19,* 606–614. doi:10.1037/h0057610

*Burgette, J. E., & Magun-Jackson, S. (2008). Freshman orientation, persistence, and achievement: A longitudinal analysis. *Journal of College Student Retention: Research, Theory and Practice, 10,* 235–263. doi:10.2190/CS.10.3.a

*Burke, P. J. (1989). Gender identity, sex, and school performance. *Social Psychology Quarterly, 52,* 159–169. doi:10.2307/2786915

*Bursik, K., & Martin, T. A. (2006). Ego development and adolescent academic achievement. *Journal of Research on Adolescence, 16,* 1–18. doi:10.1111/j.1532-7795.2006.00116.x

*Burts, D. C., Hart, C. H., Charlesworth, R., DeWolf, D. M., Ray, J., Manuel, K., & Fleege, P. O. (1993). Developmental appropriateness of kindergarten programs and academic outcomes in first grade. *Journal of Research in Childhood Education, 8,* 23–31. doi:10.1080/02568549309594852

*Busemann, A. A., & Harders, G. G. (1932). Die Wirkung väterlicher Erwerbslosigkeit auf die Schulleistungen der Kinder [The effect of paternal unemployment upon the schoolwork of children]. *Zeitschrift für Kinderforschung, 40,* 89–100.

*Büyüköztürk, S. (2004). Predictors of academic achievement for elementary teacher education students in Turkey. *International Journal of Educational Reform, 13,* 388–402.

*Byrd, C. M., & Chavous, T. M. (2009). Racial identity and academic achievement in the neighborhood context: A multilevel analysis. *Journal of Youth and Adolescence, 38,* 544–559. doi:10.1007/s10964-008-9381-9

*Byrns, R. (1930). Concerning college grades. *School & Society, 31,* 684–686.

*Calafiore, P., & Damianov, D. S. (2011). The effect of time spent online on student achievement in online economics and finance courses. *Journal of Economic Education, 42,* 209–223. doi:10.1080/00220485.2011.581934

*Call, G., Beer, J., & Beer, J. (1994). General and test anxiety, shyness, and grade point average of elementary school children of divorced and nondivorced parents. *Psychological Reports, 74,* 512–514. doi:10.2466/pr0.1994.74.2.512

Campbell, D. W., & Eaton, W. O. (1999). Sex differences in the activity level of infants. *Infant and Child Development, 8,* 1–17. doi:10.1002/(SICI)1522-7219(199903)8:1<1::AID-ICD186>3.0.CO;2-O

*Cantwell, R., Archer, J., & Bourke, S. (2001). A comparison of the academic experiences and achievement of university students entering by traditional and non-traditional means. *Assessment & Evaluation in Higher Education, 26,* 221–234. doi:10.1080/02602930120052387

Cappon, P., & Canadian Council on Learning. (2011). *Exploring the "boy crisis" in education.* Ottawa, Ontario, Canada: Canadian Council on Learning.

*Caso-Niebla, J., & Hernández-Guzmán, L. (2007). Variables que inciden en el rendimiento académico de adolescentes Mexicanos [Variables that influence academic achievement in Mexican adolescents]. *Revista Latinoamericana de Psicología, 39,* 487–501.

*Catsambis, S. (1994). The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high school. *Sociology of Education, 67,* 199–215. doi:10.2307/2112791

*Cauce, A. M. (1987). School and peer competence in early adolescence: A test of domain-specific self-perceived competence. *Developmental Psychology, 23,* 287–291. doi:10.1037/0012-1649.23.2.287

Chalabaev, A., Sarrazin, P., Trouilloud, D., & Jussim, L. (2009). Can sex undifferentiated teacher expectations mask an influence of sex stereotypes? Alternative forms of sex bias in teacher expectations. *Journal of Applied Social Psychology, 39,* 2469–2498. doi:10.1111/j.1559-1816.2009.00534.x

*Chang, C., & Chen, L. (1977). Sex differences in elementary school achievement as related to sex differences of teachers. *Bulletin of Educational Psychology, 10,* 21–33.

*Chen, J. A., & Pajares, F. (2010). Implicit theories of ability of Grade 6 science students: Relation to epistemological beliefs and academic motivation and achievement in science. *Contemporary Educational Psychology, 35,* 75–87. doi:10.1016/j.cedpsych.2009.10.003

*Chen, M., & Ehrenberg, T. (1993). Test scores, homework, aspirations and teachers' grades. *Studies in Educational Evaluation, 19,* 403–419. doi:10.1016/S0191-491X(10)80005-7

*Cheung, C. S., & McBride-Chang, C. (2008). Relations of perceived maternal parenting style, practices, and learning motivation to academic competence in Chinese children. *Merrill-Palmer Quarterly, 54,* 1–22. doi:10.1353/mpq.2008.0011

*Chittum, K. R. (1996). The relationship between physical attractiveness and psychosocial identity development, first semester GPA and admissions index scores for freshmen. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 58*(2), 0382.

*Cicirelli, V. G. (1977). Children's school grades and sibling structure. *Psychological Reports, 41,* 1055–1058. doi:10.2466/pr0.1977.41.3f.1055

*Coates, T. J., & Southern, M. L. (1972). Differential educational aspiration levels of men and women undergraduate students. *Journal of Psychology: Interdisciplinary and Applied, 81,* 125–128. doi:10.1080/00223980.1972.9923798

*Cogan, M. F. (2010). Exploring academic outcomes of homeschooled students. *Journal of College Admission, 208,* 18–25.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Craddick, R. A. (1966). Effect of season of birth on achievement of college students. *Psychological Reports, 18,* 329–330. doi:10.2466/pr0.1966.18.1.329

*Cruickshank, D. R., Kennedy, J. J., & Kapel, D. E. (1980). A case history of differential grading: Do teacher education majors really receive higher grades? *Journal of Teacher Education, 31*(4), 43–47. doi:10.1177/002248718003100411

*Daly, D. D. (2009). The relationship between college-level learning in high school and post-secondary academic success. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 70*(5), 1490.

*Dambrot, F. H., Silling, S. M., & Zook, A. (1988). Psychology of computer use: II. Sex differences in prediction of course grades in a computer language course. *Perceptual and Motor Skills, 66,* 627–636. doi:10.2466/pms.1988.66.2.627

*Daubman, K. A., Heatherington, L., & Ahn, A. (1992). Gender and the self-presentation of academic achievement. *Sex Roles, 27,* 187–204. doi:10.1007/BF00290017

*Davidson, C. W., & Haffey, P. (1979). Relationship between achievement in high school biology and type of science course completed, sex, and intelligence quotients of students. *Southern Journal of Educational Research, 13,* 133–138.

*Day, S. K. (1999). Psychological impact of attributional style and locus of control on college adjustment and academic success. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 60*(3), 0646.

*Daymont, T., & Blau, G. (2008). Student performance in online and traditional sections of an undergraduate management course. *Journal of Behavioral and Applied Management, 9,* 275–294.

*DeBerard, M. S., Spielmans, G. I., & Julka, D. L. (2004). Predictors of academic achievement and retention among college freshmen: A longitudinal study. *College Student Journal, 38,* 66–80.

*DeCoster, D. A. (1979). The effects of residence hall room visitation upon academic achievement for college students. *Journal of College Student Personnel, 20,* 520–524.

*Demirbas, O. O., & Demirkan, H. (2007). Learning styles of design students and the relationship of academic performance and gender in design education. *Learning and Instruction, 17,* 345–359. doi:10.1016/j.learninstruc.2007.02.007

*Deslandes, R., Bouchard, P., & St-Amant, J.-C. (1998). Family variables as predictors of school achievement: Sex differences in Quebec adolescents. *Canadian Journal of Education, 23,* 390–404. doi:10.2307/1585754

*Desler, M., & North, G. (1978). The impact of overassignment on grade point averages of first-time freshmen. *Journal of College and University Student Housing, 8,* 18–22.

*Dewaele, J.-M. (2007). Predicting language learners' grades in the L1, L2, L3 and L4: The effect of some psychological and sociocognitive variables. *International Journal of Multilingualism, 4,* 169–197. doi:10.2167/ijm080.0

*de Wolf, V. A. (1981). High school mathematics preparation and sex differences in quantitative abilities. *Psychology of Women Quarterly, 5,* 555–567. doi:10.1111/j.1471-6402.1981.tb00594.x

*DiLorenzo, M. L. (2009). Race-related factors in academic achievement: An examination of racial socialization and racial identity in African Americans and Latino college students. *Dissertation Abstracts International: Section B. Sciences and Engineering, 70*(5), 3219.

*Ding, C. S. (2008). Variations in academic performance trajectories during high school transition: Exploring change profiles via multidimensional scaling growth profile analysis. *Educational Research and Evaluation, 14,* 305–319. doi:10.1080/13803610802249357

*Ding, C. S., Song, K., & Richardson, L. I. (2006). Do mathematical gender differences continue? A longitudinal study of gender difference and excellence in mathematics performance in the U.S. *Educational Studies, 40,* 279–295. doi:10.1080/00131940701301952

*Dubey, R. S. (1982). Persistence, sex-differences and educational performance. *Asian Journal of Psychology & Education, 9*(3), 24–29.

*Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98,* 198–208. doi:10.1037/0022-0663.98.1.198

*Dunham, R. B. (1973). Achievement motivation as predictive of academic performance: A multivariate analysis. *Journal of Educational Research, 67,* 70–72.

*Dunkake, I., Kiechle, T., Klein, M., & Rosar, U. (2012). Schöne schüler, schöne noten? Eine empirische untersuchung zum einfluss der physischen attraktivität von schülern auf die notenvergabe durch das lehrpersonal [Good looks, good grades? An empirical analysis of the influence of students' physical attractiveness on grading by teachers]. *Zeitschrift für Soziologie, 41,* 142–161.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41,* 1040–1048. doi:10.1037/0003-066X.41.10.1040

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco, CA: Freeman.

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues, 46,* 183–201. doi:10.1111/j.1540-4560.1990.tb01929.x

*Edwards, R. P., & Thacker, K. (1979). The relationship of birth-order, gender, and sibling gender in the two-child family to grade point average in college. *Adolescence, 14,* 111–114.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634. doi:10.1136/bmj.315.7109.629

*Elliott, R., & Strenta, A. C. (1988). Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. *Journal of Educational Measurement, 25,* 333–347. doi:10.1111/j.1745-3984.1988.tb00312.x

*Elmore, P. B., & Vasu, E. S. (1980). Relationship between selected variables and statistics achievement: Building a theoretical model. *Journal of Educational Psychology, 72,* 457–467. doi:10.1037/0022-0663.72.4.457

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin, 132,* 33–72. doi:10.1037/0033-2909.132.1.33

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136,* 103–127. doi:10.1037/a0018053

*Erkman, F., Caner, A., Sart, Z. H., Börkan, B., & Şahan, K. (2010). Influence of perceived teacher acceptance, self-concept, and school attitude on the academic achievement of school-age children in Turkey. *Cross-Cultural Research, 44,* 295–309. doi:10.1177/1069397110366670

*Erkut, S. (1983). Exploring sex differences in expectancy, attribution, and academic achievement. *Sex Roles, 9,* 217–231. doi:10.1007/BF00289625

*Farkas, G., Sheehan, D., & Grobe, R. P. (1990). Coursework mastery and school success: Gender, ethnicity, and poverty groups within an urban school district. *American Educational Research Journal, 27,* 807–827. doi:10.3102/00028312027004807

*Farmer, T. W., Irvin, M. J., Thompson, J. H., Hutchins, B. C., & Leung, M.-C. (2006). School adjustment and the academic success of rural African American early adolescents in the deep south. *Journal of Research in Rural Education, 21*(3), 1–14.

*Fayowski, V., & MacMillan, P. D. (2008). An evaluation of the supplemental instruction programme in a first year calculus course. *International Journal of Mathematical Education in Science and Technology, 39,* 843–855. doi:10.1080/00207390802054433

*Feinberg, L. B., & Halperin, S. (1978). Affective and cognitive correlates of course performance in introductory statistics. *Journal of Experimental Education, 46*(4), 11–18.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43,* 95–103. doi:10.1037/0003-066X.43.2.95

*Felson, R. B. (1980). Physical attractiveness, grades and teachers' attributions of ability. *Representative Research in Social Psychology, 11,* 64–71.

*Fischbein, S. (1990). Biosocial influences on sex differences for ability and achievement test results as well as marks at school. *Intelligence, 14,* 127–139. doi:10.1016/0160-2896(90)90018-O

*Fischl, D., & Sagy, S. (2009). Factors related to students' achievements: Comparing Israeli Bedouin and Jewish students in college education. *Intercultural Education, 20,* 345–358. doi:10.1080/14675980903351979

*Flexer, B. K. (1984). Predicting eighth-grade algebra achievement. *Journal for Research in Mathematics Education, 15,* 352–360. doi:10.2307/748425

*Frailey, L. E., & Crain, C. M. (1914). Correlation of excellence in different school subjects based on a study of school grades. *Journal of Educational Psychology, 5,* 141–154. doi:10.1037/h0072639

*Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics—A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education, 22,* 497–514. doi:10.1007/BF03173468

*Friedrichsen, J. E. (1997). Self-concept and self-esteem development in the context of adolescence and gender. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 58*(12), 4565.

*Friend, C. A. (2009). The influences of parental racial socialization on the academic achievement of African American children: A cultural-ecological approach. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 70*(5), 1553.

*Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2002). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences, 14,* 49–66. doi:10.1016/j.lindif.2003.08.002

*Gadzella, B. M., Cochran, S. W., Parham, L., & Fournet, G. P. (1976). Accuracy and differences among students in their predictions of semester achievement. *Journal of Educational Research, 70,* 75–81.

*Gadzella, B. M., Williamson, J. D., & Ginther, D. W. (1985). Correlations of self-concept with locus of control and academic performance. *Perceptual and Motor Skills, 61,* 639–645. doi:10.2466/pms.1985.61.2.639

Garson, G. D. (2013). Preparing to analyze multilevel data. In G. D. Garson (Ed.), *Hierarchical linear modeling: Guide and applications* (pp. 27–54). Los Angeles, CA: Sage.

*Gerberich, S. G., Gibson, R. W., Fife, D., Mandel, J. S., Aeppli, D., Le, C. T., . . . Matross, R. (1997). Effects of brain injury on college academic performance. *Neuroepidemiology, 16,* 1–14. doi:10.1159/000109665

*Giancarlo, C. A., & Facione, P. A. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *Journal of General Education, 50,* 29–55. doi:10.1353/jge.2001.0004

*Gifford, D. D., Briceno-Perriott, J., & Mianzo, F. (2006). Locus of control: Academic achievement and retention in a sample of university first-year students. *Journal of College Admission, 191,* 18–25.

*Gillock, K. L., & Reyes, O. (1999). Stress, support, and academic performance of urban, low-income, Mexican-American adolescents. *Journal of Youth and Adolescence, 28,* 259–282. doi:10.1023/A:1021657516275

*Glass, J. C., & Garrett, M. S. (1995). Student participation in a college orientation course, retention, and grade point average. *Community College Journal of Research and Practice, 19,* 117–132. doi:10.1080/1066892950190203

*Goetz, T., Frenzel, A. C., Hall, N. C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology, 33,* 9–33. doi:10.1016/j.cedpsych.2006.12.002

Goldberg, W. A., Prause, J. A., Lucas-Thompson, R., & Himsel, A. (2008). Maternal employment and children's achievement in context: A meta-analysis of four decades of research. *Psychological Bulletin, 134,* 77–108. doi:10.1037/0033-2909.134.1.77

Goodenough, F. L. (1954). The measurement of mental growth in children. In L. Carmichael (Ed.), *Manual of child psychology* (2nd ed., pp. 459–491). New York, NY: Wiley.

*Goodman, A., & Koupil, I. (2010). The effect of school performance upon marriage and long-term reproductive success in 10,000 Swedish males and females born 1915–1929. *Evolution and Human Behavior, 31,* 425–435. doi:10.1016/j.evolhumbehav.2010.06.002

*Goodman, S. B., & Cirka, C. C. (2009). Efficacy and anxiety: An examination of writing attitudes in a first-year seminar. *Journal on Excellence in College Teaching, 20*(3), 5–28.

*Goodstein, L. D., Crites, J. O., & Heilbrun, A. B. J. (1963). Personality correlates of academic adjustment. *Psychological Reports, 12,* 175–196. doi:10.2466/pr0.1963.12.1.175

*Graham, L. D. (1991). Predicting academic success of students in a master of business administration program. *Educational and Psychological Measurement, 51,* 721–727. doi:10.1177/0013164491513023

*Grave, B. S. (2011). The effect of student time allocation on academic achievement. *Education Economics, 19,* 291–310. doi:10.1080/09645292.2011.585794

*Gupta, S., Harris, D. E., Carrier, N. M., & Caron, P. (2006). Predictors of student success in entry-level undergraduate mathematics courses. *College Student Journal, 40,* 97–108.

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology Press.

Halpern, D. F., Straight, C. A., & Stephenson, C. L. (2011). Beliefs about cognitive gender differences: Accurate for direction, underestimated for size. *Sex Roles, 64,* 336–347. doi:10.1007/s11199-010-9891-2

*Ham, B. D. (2004). The effects of divorce and remarriage on the academic achievement of high school seniors. *Journal of Divorce & Remarriage, 42,* 159–178. doi:10.1300/J087v42n01_08

*Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72,* 625–638. doi:10.1111/1467-8624.00301

*Hancock, T. (1999). The gender difference: Validity of standardized admission tests in predicting MBA performance. *Journal of Education for Business, 75,* 91–93. doi:10.1080/08832329909598996

*Hanna, G. S., & Sonnenschein, J. L. (1985). Relative validity of the Orleans-Hanna Algebra Prognosis Test in the prediction of girls' and boys' grades in first-year algebra. *Educational and Psychological Measurement, 45,* 361–367. doi:10.1177/001316448504500221

*Harris, E. W., Tanner, J. R., & Knouse, S. B. (1996). Employment of recent university business graduates: Do age, gender, and minority status make a difference? *Journal of Employment Counseling, 33,* 121–129. doi:10.1002/j.2161-1920.1996.tb00444.x

*Harrison, C. S. (1996). Relationships between grades in music theory for nonmusic majors and selected background variables. *Journal of Research in Music Education, 44,* 341–352. doi:10.2307/3345446

Hartley, B. L., & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development, 84,* 1716–1733. doi:10.1111/cdev.12079

*Healy, C. C., Tullier, M., & Mourton, D. M. (1990). My vocational situation: Its relation to concurrent career and future academic benchmarks. *Measurement and Evaluation in Counseling and Development, 23,* 100–107.

*Heatherington, L., Daubman, K. A., Bates, C., Ahn, A., Brown, H., & Preston, C. (1993). Two investigations of "female modesty" in achievement situations. *Sex Roles, 29,* 739–754. doi:10.1007/BF00289215

*Heatherington, L., Townsend, L. S., & Burroughs, D. P. (2001). "How'd you do on that test?": The effects of gender and self-presentation of achievement to vulnerable men. *Sex Roles, 45,* 161–177. doi:10.1023/A:1013597626641

*Heaven, P. C. L., Ciarrochi, J., & Vialle, W. (2007). Conscientiousness and Eysenckian psychoticism as predictors of school grades: A one-year longitudinal study. *Personality and Individual Differences, 42,* 535–546. doi:10.1016/j.paid.2006.07.028

Hedges, L. V., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 14–50). Baltimore, MD: Johns Hopkins University Press.

Hedges, L. V., & Nowell, A. (1995, July 7). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269,* 41–45. doi:10.1126/science.7604277

*Helbig, M. (2010). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? [Are female teachers responsible for the lower educational attainment of boys?]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 62,* 93–111. doi:10.1007/s11577-010-0095-0

*Herron, E. W. (1964). Relationship of experimentally aroused achievement motivation to academic achievement anxiety. *Journal of Abnormal and Social Psychology, 69,* 690–694. doi:10.1037/h0044038

*Hewitt, B. N., & Goldman, R. D. (1975). Occam's razor slices through the myth that college women overachieve. *Journal of Educational Psychology, 67,* 325–330. doi:10.1037/h0077010

*Hildenbrand, K. J. (2005). An examination of college student athletes' academic achievement. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 66*(12), 4308.

*Hogan, M. J., Parker, J. D. A., Wiener, J., Watters, C., Wood, L. M., & Oke, A. (2010). Academic success in adolescence: Relationships among verbal IQ, social support and emotional intelligence. *Australian Journal of Psychology, 62,* 30–41. doi:10.1080/00049530903312881

*Horvath, J., Beaudin, B. Q., & Wright, S. P. (1992). Persisting in the introductory economics course: An exploration of gender differences. *Journal of Economic Education, 23,* 101–108. doi:10.2307/1183251

*Hosseini, A. A. (1975). Comparison of successful and unsuccessful students at Pahlavi University. *Psychological Reports, 36,* 15–20. doi:10.2466/pr0.1975.36.1.15

*House, J. D., & Keeley, E. J. (1995). Gender bias in prediction of graduate grade performance from Miller Analogies Test scores. *Journal of Psychology: Interdisciplinary and Applied, 129,* 353–355. doi:10.1080/00223980.1995.9914972

*Houston, L. N. (1987). The predictive validity of a study habits inventory for first semester undergraduates. *Educational and Psychological Measurement, 47,* 1025–1030. doi:10.1177/0013164487474018

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Hove, England: Routledge.

*Hudy, G. T. (2006). An analysis of motivational factors related to academic success and persistence for university students. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 67*(11), 4094.

*Hughey, A. W. (1995). Observed differences in Graduate Record Examination scores and mean undergraduate grade point averages by gender and race among students admitted to a master's degree program in college student affairs. *Psychological Reports, 77,* 1315–1321. doi:10.2466/pr0.1995.77.3f.1315

*Hunley, S. A., Evans, J. H., Delgado-Hachey, M., Krise, J., Rich, T., & Schell, C. (2005). Adolescent computer use and academic achievement. *Adolescence, 40,* 307–318.

*Huysamen, G. K., & Roozendaal, L. A. (1999). Curricular choice and the differential prediction of the tertiary-academic performance of men and women. *South African Journal of Psychology, 29,* 87–93. doi:10.1177/008124639902900205

Hyde, J. S., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139–155. doi:10.1037/0033-2909.107.2.139

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008, July 25). Gender similarities characterize math performance. *Science, 321,* 494–495. doi:10.1126/science.1160364

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104,* 53–69. doi:10.1037/0033-2909.104.1.53

*Imms, W. D. (2000). Boys' rates of participation and academic achievement in international baccalaureate art and design education. *Visual Arts Research, 26,* 53–74.

*Ismail, N. A., & Othman, A. (2006). Comparing university academic performances of HSC students at the three art-based faculties. *International Education Journal, 7,* 668–675.

*Jacobowitz, T. (1983). Relationship of sex, achievement, and science self-concept to the science career preferences of Black students. *Journal of Research in Science Teaching, 20,* 621–628. doi:10.1002/tea.3660200703

*Jansen, E. P. W. A., & Bruinsma, M. (2005). Explaining achievement in higher education. *Educational Research and Evaluation, 11,* 235–252. doi:10.1080/13803610500101173

*Johnson, M., & Kuennen, E. (2006). Basic math skills and performance in an introductory statistics course. *Journal of Statistics Education, 14*(2). Retrieved from http://www.amstat.org/publications/jse/v14n2/johnson.html

*Johnston, M. W. (1999). The relationship between locus of control and academic achievement for adults (gender, race). *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 60*(4), 1010.

*Jones, P. S. (2010). The impact of four-year participation in music and/or athletic activities in South Dakota public high schools on GPA and ACT scores. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 71*(8), 2860.

Jussim, L., Robustelli, S., & Cain, T. (2009). Teacher expectations and self-fulfilling prophecies. In A. Wigfield & K. Wentzel (Eds.), *Handbook of motivation at school* (pp. 349–380). Mahwah, NJ: Erlbaum.

*Kaczmarek, M., & Franco, J. N. (1986). Sex differences in prediction of academic performance by the Graduate Record Exam. *Psychological Reports, 59,* 1197–1198. doi:10.2466/pr0.1986.59.3.1197

Kalaian, H. A., & Raudenbush, S. W. (1996). A multivariate mixed linear model for meta-analysis. *Psychological Methods, 1,* 227–235. doi:10.1037/1082-989X.1.3.227

*Keiller, S. W. (1997). Type A behavior pattern and locus of control as predictors of college academic performance. *Dissertation Abstracts International: Section B. Sciences and Engineering, 58*(5), 2737.

*Keith, T. Z., & Benson, M. J. (1992). Effects of manipulable influences on high school grades across five ethnic groups. *Journal of Educational Research, 86,* 85–93. doi:10.1080/00220671.1992.9941144

*Keller, D., Crouse, J., & Trusheim, D. (1993). Relationships among gender differences in freshman course grades and course characteristics. *Journal of Educational Psychology, 85,* 702–709. doi:10.1037/0022-0663.85.4.702

*Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M., & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology, 42,* 11–26. doi:10.1037/0012-1649.42.1.11

*Khan, S., Haynes, L., Armstrong, A., & Rohner, R. P. (2010). Perceived teacher acceptance, parental acceptance, academic achievement, and school conduct of middle school students in the Mississippi Delta region of the United States. *Cross-Cultural Research, 44,* 283–294. doi:10.1177/1069397110368030

Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105,* 198–214. doi:10.1037/0033-2909.105.2.198

*King, D. B., & Joshi, S. (2008). Gender differences in the use and effectiveness of personal response devices. *Journal of Science Education and Technology, 17,* 544–552. doi:10.1007/s10956-008-9121-7

*Kitsantas, A., & Zimmerman, B. J. (2009). College students' homework and academic achievement: The mediating role of self-regulatory beliefs. *Metacognition and Learning, 4,* 97–110. doi:10.1007/s11409-008-9028-y

*Klugh, H. E., & Bierley, R. (1959). The School and College Ability Test and high school grades as predictors of college achievement. *Educational and Psychological Measurement, 19,* 625–626. doi:10.1177/001316445901900415

*Koenig, J. A., Sireci, S. G., & Wiley, A. (1998). Evaluating the predictive validity of MCAT scores across diverse applicant groups. *Academic Medicine, 73,* 1095–1106. doi:10.1097/00001888-199810000-00021

*Kokkelenberg, E. C., & Sinha, E. (2010). Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students. *Economics of Education Review, 29,* 935–946. doi:10.1016/j.econedurev.2010.06.016

*Kollárik, K. (1991). Rozdiely medzi chlapcami a dievčatami v úrovni pripravenosti na školu a v učebných výsledkoch [Differences between boys and girls in their school readiness level and school achievements]. *Psychológia A Patopsychológia Diet'at'a, 26,* 113–124.

*Kost, L. E., Pollock, S. J., & Finkelstein, N. D. (2009). Characterizing the gender gap in introductory physics. *Physical Review Special Topics—Physics Education Research, 5*(1), 1–14. doi:10.1103/PhysRevSTPER.5.010101

*Kucerova, L. (1975). Relationship between intellectual abilities and educational success of girls and boys during their first school years. *Psychológia A Patopsychológia Diet'at'a, 10,* 71–78.

Kuncel, N. R., Henzltet, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127,* 162–181. doi:10.1037/0033-2909.127.1.162

*Kurdek, L. A., & Sinclair, R. J. (1988). Relation of eighth graders' family structure, gender, and family environment with academic performance and school behavior. *Journal of Educational Psychology, 80,* 90–94. doi:10.1037/0022-0663.80.1.90

*Lagerström, M., Bremme, K., Eneroth, P., & Magnusson, D. (1991). Sex-related differences in school and IQ performance for children with low birth weight at ages 10 and 13. *Journal of Special Education, 25,* 261–270. doi:10.1177/002246699102500209

*Lee, F. H., & Nemzek, C. L. (1941). Relation between certain physical defects and school achievement. *Journal of Social Psychology, 13,* 385–394. doi:10.1080/00224545.1941.9714086

*Lehn, T., Vladovic, R., & Michael, W. B. (1980). The short-term predictive validity of a standardized reading test and of scales reflecting six dimensions of academic self-concept relative to selected high school achievement criteria for four ethnic groups. *Educational and Psychological Measurement, 40,* 1017–1031. doi:10.1177/001316448004000429

*Lehre, A.-C., Hansen, A., & Laake, P. (2009). Gender and the 2003 quality reform in higher education in Norway. *Higher Education, 58,* 585–597. doi:10.1007/s10734-009-9211-3

*Lekholm, A. K., & Cliffordson, C. (2008). Discrepancies between school grades and test scores at individual and school level: Effects of gender and family background. *Educational Research and Evaluation, 14,* 181–199. doi:10.1080/13803610801956663

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136,* 1123–1135. doi:10.1037/a0021276

*Lindley, L. D., & Borgen, F. H. (2002). Generalized self-efficacy, Holland theme self-efficacy, and academic performance. *Journal of Career Assessment, 10,* 301–314. doi:10.1177/10672702010003002

*Lindsay, C. A., & Althouse, R. (1969). Comparative validities of the Strong Vocational Interest Blank Academic Achievement Scale and the College Student Questionnaire Motivation for Grades Scale. *Educational and Psychological Measurement, 29,* 489–493. doi:10.1177/001316446902900229

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

*Llabre, M. M., & Suarez, E. (1985). Predicting math anxiety and course performance in college women and men. *Journal of Counseling Psychology, 32,* 283–287. doi:10.1037/0022-0167.32.2.283

*Lloyd, J. E. V., Walsh, J., & Yailagh, M. S. (2005). Sex differences in performance attributions, self-efficacy, and achievement in mathematics: If I'm so smart, why don't I know it? *Canadian Journal of Education, 28,* 384–408. doi:10.2307/4126476

*Long, J. F., Monoi, S., Harper, B., Knoblauch, D., & Murphy, P. K. (2007). Academic motivation and achievement among urban adolescents. *Urban Education, 42,* 196–222. doi:10.1177/0042085907300447

*Lumme, K., & Lehto, J. E. (2002). Sixth grade pupils' phonological processing and school achievement in a second and native language. *Scandinavian Journal of Educational Research, 46,* 207–217. doi:10.1080/00313830220142209

*Lunneborg, P. W. (1977). Longitudinal criteria of predicted college achievement. *Measurement & Evaluation in Guidance, 9,* 212–213.

*Lutz, A., & Crist, S. (2009). Why do bilingual boys get better grades in English-only America? The impacts of gender, language and family interaction on academic achievement of Latino/a children of immigrants.

Ethnic and Racial Studies, 32, 346–368. doi:10.1080/01419870801943647

Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *Trames, 13,* 3–13. doi:10.3176/tr.2009.1.01

*Maqsud, M. (1993). Relationships of some personality variables to academic attainment of secondary school pupils. *Educational Psychology, 13,* 11–18. doi:10.1080/0144341930130102

Marsh, H. W., Bornmann, L., Mutz, R., Daniel, H.-D., & O'Mara, A. (2009). Gender effects in the peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research, 79,* 1290–1326. doi:10.3102/0034654309334143

*Matthews, D. B. (1991). The effects of learning style on grades of first-year college students. *Research in Higher Education, 32,* 253–268. doi:10.1007/BF00992891

*Mau, W.-C., & Lynn, R. (2001). Gender differences on the Scholastic Aptitude Test, the American College Test and college grades. *Educational Psychology, 21,* 133–136. doi:10.1080/01443410020043832

*McCandless, B. R., Roberts, A., & Starnes, T. (1972). Teachers' marks, achievement test scores, and aptitude relations with respect to social class, race, and sex. *Journal of Educational Psychology, 63,* 153–159. doi:10.1037/h0032646

*McCornack, R. L., & McLeod, M. M. (1988). Gender bias in the prediction of college course performance. *Journal of Educational Measurement, 25,* 321–331. doi:10.1111/j.1745-3984.1988.tb00311.x

*McDonald, J. F., & McPherson, M. S. (1975). High school type, sex, and socio-economic factors as predictors of the academic achievement of university students. *Educational and Psychological Measurement, 35,* 929–933. doi:10.1177/001316447503500421

Mead, S. (2006). *The evidence suggests otherwise: The truth about boys and girls*. Retrieved from http://www.educationsector.org/usr_doc/ESO_BoysAndGirls.pdf

*Mickelson, R. A., & Greene, A. D. (2006). Connecting pieces of the puzzle: Gender differences in Black middle school students' achievement. *Journal of Negro Education, 75,* 34–48.

*Miller, C. D., Finley, J., & McKinley, D. L. (1990). Learning approaches and motives: Male and female differences and implications for learning assistance programs. *Journal of College Student Development, 31,* 147–154.

*Mills, C., Heyworth, J., Rosenwax, L., Carr, S., & Rosenberg, M. (2009). Factors associated with the academic success of first year health science students. *Advances in Health Sciences Education, 14,* 205–217. doi:10.1007/s10459-008-9103-9

*Morganson, V. J., Jones, M. P., & Major, D. A. (2010). Understanding women's underrepresentation in science, technology, engineering, and mathematics: The role of social coping. *Career Development Quarterly, 59,* 169–179. doi:10.1002/j.2161-0045.2010.tb00060.x

*Mpofu, E., D'Amico, M., & Cleghorn, A. (1996). Time management practices in an African culture: Correlates with college academic grades. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 28,* 102–112. doi:10.1037/0008-400X.28.2.102

*Mullola, S., Jokela, M., Ravaja, N., Lipsanen, J., Hintsanen, M., Alatupa, S., & Keltikangas-Järvinen, L. (2011). Associations of student temperament and educational competence with academic achievement: The role of teacher age and teacher and student gender. *Teaching and Teacher Education, 27,* 942–951. doi:10.1016/j.tate.2011.03.005

*Neighbors, B., Forehand, R., & Armistead, L. (1992). Is parental divorce a critical stressor for young adolescents? Grade point average as a case in point. *Adolescence, 27,* 639–646.

*Nelson, D. D. (1969). A study of school achievement among adolescent children with working and nonworking mothers. *Journal of Educational Research, 62,* 456–458.

*Nguyen, N. T., Allen, L. C., & Fraccastoro, K. (2005). Personality predicts academic performance: Exploring the moderating role of gender. *Journal of Higher Education Policy and Management, 27,* 105–117. doi:10.1080/13600800500046313

*Nicpon, M. F., Huser, L., Blanks, E. H., Sollenberger, S., Befort, C., & Kurpius, S. E. R. (2006). The relationship of loneliness and social support with college freshmen's academic performance and persistence. *Journal of College Student Retention: Research, Theory & Practice, 8,* 345–358. doi:10.2190/A465-356M-7652-783R

*Nishina, A., Juvonen, J., & Witkow, M. R. (2005). Sticks and stones may break my bones, but names will make me feel sick: The psychosocial, somatic, and scholastic consequences of peer harassment. *Journal of Clinical Child and Adolescent Psychology, 34,* 37–48. doi:10.1207/s15374424jccp3401_4

Nordvik, H., & Amponsah, B. (1998). Gender differences in spatial abilities and spatial activity among university students in an egalitarian educational system. *Sex Roles, 38,* 1009–1023. doi:10.1023/A:1018878610405

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., . . . Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *PNAS: Proceedings of the National Academy of Sciences, USA, 106,* 10593–10597. doi:10.1073/pnas.0809921106

Nowell, A., & Hedges, L. V. (1998). Trends in gender differences in academic achievement from 1960 to 1994: An analysis of differences in mean, variance, and extreme scores. *Sex Roles, 39,* 21–43. doi:10.1023/A:1018873615316

*Odell, P. M., & Schumacher, P. (1998). Attitudes towards mathematics and predictors of college mathematics grades: Gender differences in a 4-year business college. *Journal of Education for Business, 74,* 34–38. doi:10.1080/08832329809601658

*Öhlund, L. S., & Ericsson, K. B. (1994). Elementary school achievement and absence due to illness. *Journal of Genetic Psychology, 155,* 409–421. doi:10.1080/00221325.1994.9914791

*Olds, D. E., & Shaver, P. (1980). Masculinity, femininity, academic performance, and health: Further evidence concerning the androgyny controversy. *Journal of Personality, 48,* 323–341. doi:10.1111/j.1467-6494.1980.tb00837.x

*O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal, 44,* 161–196. doi:10.3102/0002831206298171

*Paolillo, J. G. P. (1982). The predictive validity of selected admissions variables relative to grade point average earned in a master of business administration program. *Educational and Psychological Measurement, 42,* 1163–1167. doi:10.1177/001316448204200423

Patrick, H., Ryan, A. M., & Pintrich, P. R. (1999). The differential impact of extrinsic and mastery goal orientations on males' and females' self-regulated learning. *Learning and Individual Differences, 11,* 153–171. doi:10.1016/S1041-6080(00)80003-5

*Payne, D. A., Rapley, F. E., & Wells, R. A. (1973). Application of a biographical data inventory to estimate college academic achievement. *Measurement & Evaluation in Guidance, 6,* 152–156.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). Fort Worth, TX: Harcourt Brace.

*Pedrini, B. C., & Pedrini, D. T. (1978). Grade point evaluation of an experimental program for disadvantaged college freshmen. *College Student Journal, 11,* 318–323.

*Perrault, Y. L. (1976). Sex, motivation, extraversion and neuroticism as multiple predictors and moderators with predicting two grade nine attainment criteria. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 68*(7), 2812.

*Phillips, B. N. (1962). Sex, social class, and anxiety as sources of variation in school achievement. *Journal of Educational Psychology, 53,* 316–322. doi:10.1037/h0048211

*Pomerantz, E. M., Altermatt, E. R., & Saxon, J. L. (2002). Making the grade but feeling distressed: Gender differences in academic performance and internal distress. *Journal of Educational Psychology, 94,* 396–404. doi:10.1037/0022-0663.94.2.396

*Post, T. R., Medhanie, A., Harwell, M., Norman, K. W., Dupuis, D. N., Muchlinski, T., . . . Monson, D. (2010). The impact of prior mathematics achievement on the relationship between high school mathematics curricula and postsecondary mathematics performance, course-taking, and persistence. *Journal for Research in Mathematics Education, 41,* 274–308.

*Preckel, F., Goetz, T., Pekrun, R., & Kleine, M. (2008). Gender differences in gifted and average-ability students: Comparing girls' and boys' achievement, self-concept, interest, and motivation in mathematics. *Gifted Child Quarterly, 52,* 146–159. doi:10.1177/0016986208315834

*Preiss, M., & Fráňová, L. (2006). Depressive symptoms, academic achievement, and intelligence. *Studia Psychologica, 48,* 57–67.

*Preszler, R. W. (2009). Replacing lecture with peer-led workshops improves student learning. *CBE: Life Sciences Education, 8,* 182–192. doi:10.1187/cbe.09-01-0002

*Pulvino, C. J., & Hansen, J. C. (1972). Relevance of "needs" and "press" to anxiety, alienation and GPA. *Journal of Experimental Education, 40,* 70–75.

*Quirk, K. J., Keith, T. Z., & Quirk, J. T. (2001). Employment during high school and student achievement: Longitudinal analysis of national data. *Journal of Educational Research, 95,* 4–10. doi:10.1080/00220670109598778

*Ramsbottom-Lucier, M., Johnson, M. M. S., & Elam, C. L. (1995). Age and gender differences in students' preadmission qualifications and medical school performances. *Academic Medicine, 70,* 236–239. doi:10.1097/00001888-199503000-00016

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

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., Jr., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling user's manual*. Lincolnwood, IL: Scientific Software International.

Raymond, C. L., & Benbow, C. P. (1989). Educational encouragement by parents: Its relationship to precocity and gender. *Gifted Child Quarterly, 33,* 144–151. doi:10.1177/001698628903300404

*Rech, J. F. (1996). Gender differences in mathematics achievement and other variables among university students. *Journal of Research & Development in Education, 29*(2), 73–76.

*Ritchie, K. (2003). Factors affecting the transition to university. *Dissertation Abstracts International: Section B. Sciences and Engineering, 65*(1), 482.

*Rochelle, C. F., & Dotterweich, D. (2007). Student success in business statistics. *Journal of Economics and Finance Education, 6,* 19–24.

*Rogers, M. A., Theule, J., Ryan, B. A., Adams, G. R., & Keating, L. (2009). Parental involvement and children's school achievement: Evidence for mediating processes. *Canadian Journal of School Psychology, 24,* 34–57. doi:10.1177/0829573508328445

*Romine, P. G., & Quattlebaum, R. F. (1976). The Achievement Motivation Scale revisited: A cross-validation study. *Educational and Psychological Measurement, 36,* 1069–1073. doi:10.1177/001316447603600437

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (rev. ed.). Beverly Hills, CA: Sage.

*Ross, S., & Horner, B. (1949). Still more on "the physique-temperament correlation." *Journal of Heredity, 40,* 265.

*Rothstein, D. S. (2007). High school employment and youths' academic achievement. *Journal of Human Resources, 42,* 194–213.

*Ruban, L. M., & McCoach, D. B. (2005). Gender differences in explaining grades using structural equation modeling. *Review of Higher Education, 28,* 475–502. doi:10.1353/rhe.2005.0049

*Rustemeyer, R., & Fischer, N. (2005). Sex- and age-related differences in mathematics. *Psychological Reports, 97,* 183–194. doi:10.2466/pr0.97.1.183-194

Rybak Lang, R. (2011, June). A retrospective look: The "boy crisis." *The Quick and the Ed, 2011*(6). Retrieved from http://www.quickanded.com/2011/06/a-retrospective-look-the-boy-crisis.html

*Sahin, M. (2009). Correlations of students' grades, expectations, epistemological beliefs and demographics in a problem-based introductory physics course. *International Journal of Environmental and Science Education, 4,* 169–184.

*Sampson, C., & Boyer, P. G. (2001). GRE scores as predictors of minority students' success in graduate study: An argument for change. *College Student Journal, 35,* 271–279.

*Saunders, J., Davis, L., Williams, T., & Williams, J. H. (2004). Gender differences in self-perceptions and academic outcomes: A study of African American high school students. *Journal of Youth and Adolescence, 33,* 81–90. doi:10.1023/A:1027390531768

*Schaffer, B. F., Ahmadi, H., & Calkins, D. O. (1986). Academic qualifications of women with degrees in business. *Journal of Education for Business, 61,* 321–324. doi:10.1080/08832323.1986.10772738

*Schneider, L. J., & Overton, T. D. (1983). Holland personality types and academic achievement. *Journal of Counseling Psychology, 30,* 287–289. doi:10.1037/0022-0167.30.2.287

*Schulenberg, J. E., Asp, C. E., & Petersen, A. C. (1984). School from the young adolescent's perspective: A descriptive report. *Journal of Early Adolescence, 4,* 107–130. doi:10.1177/0272431684042002

*Scott, S. G. (2010). Enhancing reflection skills through learning portfolios: An empirical test. *Journal of Management Education, 34,* 430–457. doi:10.1177/1052562909351144

*Seginer, R., & Vermulst, A. (2002). Family environment, educational aspirations, and academic achievement in two cultural settings. *Journal of Cross-Cultural Psychology, 33,* 540–558. doi:10.1177/00220022102238268

*Sendelbach, W. (1975). School achievement and behavior ratings used in the validation of school entrance tests: An investigation at the end of second grade. *Psychologie in Erziehung und Unterricht, 22,* 23–31.

*Senfeld, L. (1995). Math anxiety and its relationship to selected student attitudes and beliefs. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 56*(7), 2598.

*Sexton, D. W., & Goldman, R. D. (1975). High school transcript as a set of "nonreactive" measures for predicting college success and major field. *Journal of Educational Psychology, 67,* 30–37. doi:10.1037/h0078668

*Seyfried, S. F. (1998). Academic achievement of African American preadolescents: The influence of teacher perceptions. *American Journal of Community Psychology, 26,* 381–402. doi:10.1023/A:1022107120472

*Sheard, M. (2009). Hardiness commitment, gender, and age differentiate university academic performance. *British Journal of Educational Psychology, 79,* 189–204. doi:10.1348/000709908X304406

*Sherman, J. A. (1980). Predicting mathematics grades of high school girls and boys: A further study. *Contemporary Educational Psychology, 5,* 249–255. doi:10.1016/0361-476X(80)90048-X

*Sherman, L. W., & Hofmann, R. J. (1980). Achievement as a momentary event, as a continuing state, locus of control: A clarification. *Perceptual and Motor Skills, 51,* 1159–1166. doi:10.2466/pms.1980.51.3f.1159

*Shields, N. (2001). Stress, active coping, and academic performance among persisting and nonpersisting college students. *Journal of Applied Biobehavioral Research, 6,* 65–81. doi:10.1111/j.1751-9861.2001.tb00107.x

*Shores, M. L., Smith, T. G., & Jarrell, S. L. (2009). Do individual learner variables contribute to differences in mathematics performance? *SRATE Journal, 18*(2), 1–8.

*Sibulkin, A. E., & Butler, J. S. (2008). Should college algebra be a prerequisite for taking psychology statistics? *Teaching of Psychology, 35,* 214–217. doi:10.1080/00986280802286399

*Simon, L. S. (1978). Prediction of achievement in clinical pharmacy courses. *American Journal of Pharmaceutical Education, 42,* 111–114.

*Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42,* 70–83. doi:10.1037/0012-1649.42.1.70

*Singleton, R. A. J. (2007). Collegiate alcohol consumption and academic performance. *Journal of Studies on Alcohol and Drugs, 68,* 548–555.

*Smith, T. L. (2008). The impact of residential community living learning programs on college student achievement and behavior. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 69*(8), 2975.

*Smrtnik-Vitulićc, H., & Zupančič, M. (2011). Personality traits as a predictor of academic achievement in adolescents. *Educational Studies, 37,* 127–140. doi:10.1080/03055691003729062

*Snyder, R. F. (2000). The relationship between learning styles/multiple intelligences and academic achievement of high school students. *High School Journal, 83*(2), 11–20.

*Soares, A. P., Guisande, M. A., Almeida, L. S., & Páramo, M. F. (2009). Academic achievement in first-year Portuguese college students: The role of academic preparation and learning strategies. *International Journal of Psychology, 44,* 204–212. doi:10.1080/00207590701700545

*Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality, 22,* 185–209. doi:10.1002/per.676

*Stephens, L. J., & Schaben, L. A. (2002, March). The effect of interscholastic sports participation on academic achievement of middle level school students. *NASSP Bulletin, 86*(630), 34–41. doi:10.1177/019263650208663005

*Stockard, J., Lang, D., & Wood, J. W. (1985). Academic merit, status variables, and students' grades. *Journal of Research & Development in Education, 18*(2), 12–20.

*Stupnisky, R. H., Renaud, R. D., Perry, R. P., Ruthig, J. C., Haynes, T. L., & Clifton, R. A. (2007). Comparing self-esteem and perceived control as predictors of first-year college students' academic achievement. *Social Psychology of Education, 10,* 303–330. doi:10.1007/s11218-007-9020-4

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin, 135,* 859–884. doi:10.1037/a0017364

*Sulaiman, A., & Mohezar, S. (2006). Student success factors: Identifying key predictors. *Journal of Education for Business, 81,* 328–333. doi:10.3200/JOEB.81.6.328-333

*Sullivan, A., & Voyer, D. (1998). [Gender differences in mental rotation, actual grades, and self-reported grades]. Unpublished raw data.

*Sullivan-Ham, K. (2010). Impact of participation in a dual enrollment program on first semester college GPA. *Dissertation Abstracts International: Section B. Sciences and Engineering, 71*(11), 7069.

*Sweet, P. R., & Nuttall, R. L. (1971). The effects of a tracking system on student satisfaction and achievement. *American Educational Research Journal, 8,* 511–520. doi:10.3102/00028312008003511

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

*Talento-Miller, E. (2008). Generalizability of GMAT® validity to programs outside the U.S. *International Journal of Testing, 8,* 127–142. doi:10.1080/15305050802001193

*Taube, S. R., & Taube, P. M. (1990). Pre- and post-enrollment factors associated with achievement in technical postsecondary schools. *Community/Junior College Research Quarterly of Research and Practice, 14,* 93–99. doi:10.1080/0361697900140203

*Taylor, D. J., Clay, K. C., Bramoweth, A. D., Sethi, K., & Roane, B. M. (2011). Circadian phase preference in college students: Relationships with psychological functioning and academics. *Chronobiology International, 28,* 541–547. doi:10.3109/07420528.2011.580870

*Tenenbaum, H. R., & Leaper, C. (2003). Parent–child conversations about science: The socialization of gender inequities? *Developmental Psychology, 39,* 34–47. doi:10.1037/0012-1649.39.1.34

*Thiede, W. B. (1950). Some characteristics of juniors enrolled in selected curricula at the University of Wisconsin. *Journal of Experimental Education, 19,* 1–62.

*Thomas, C. L. (1979). Relative effectiveness of high school grades for predicting college grades: Sex and ability level effects. *Journal of Negro Education, 48,* 6–13. doi:10.2307/2294611

*Thompson, J., Samiratedu, V., & Rafter, J. (1993). The effects of on-campus residence on first-time college students. *NASPA Journal, 31,* 41–47.

*Ting, S.-M. R., & Robinson, T. L. (1998). First-year academic success: A prediction combining cognitive and psychosocial variables for Caucasian and African American students. *Journal of College Student Development, 39,* 599–610.

*Tiruneh, G. (2007). Does attendance enhance political science grades? *Journal of Political Science Education, 3,* 265–276. doi:10.1080/15512160701620776

*Treiber, J. (2010). Conscientiousness, openness, and gender as academic predictive variables in high school seniors. *Dissertation Abstracts International: Section B. Sciences and Engineering, 72*(2), 1202.

*Trent, E. R. (1974). An analysis of sex differences in psychological differentiation. *Dissertation Abstracts International, 35*(5), 2416.

*Trippi, J. F., & Baker, S. B. (1989). Student and residential correlates of Black student grade performance and persistence at a predominantly White university campus. *Journal of College Student Development, 30,* 136–143.

*Truell, A. D., Zhao, J. J., Alexander, M. W., & Hill, I. B. (2006). Predicting final student performance in a graduate business program: The MBA. *Delta Pi Epsilon Journal, 48,* 144–152.

*Tulviste, T., & Rohner, R. P. (2010). Relationships between perceived teachers' and parental behavior and adolescent outcomes in Estonia. *Cross-Cultural Research, 44,* 222–238. doi:10.1177/1069397110366797

Tyre, P. (2006, January 29). Education: Boys falling behind girls in many areas. *Newsweek*. Retrieved from http://www.newsweek.com/education-boys-falling-behind-girls-many-areas-108593

*Undheim, J. O., & Nordvik, H. (1992). Socio-economic factors and sex differences in an egalitarian educational system: Academic achievement in 16-year-old Norwegian students. *Scandinavian Journal of Educational Research, 36,* 87–98. doi:10.1080/0031383920360201

*Urdan, T. C. (1997). Examining the relations among early adolescent students' goals and friends' orientation toward effort and achievement in school. *Contemporary Educational Psychology, 22,* 165–191. doi:10.1006/ceps.1997.0930

Vail, K. (2006). Is the boy crisis real? *American School Board Journal, 193,* 22–23.

*Valenzuela, A. (1993). Liberal gender role attitudes and academic achievement among Mexican-origin adolescents in two Houston inner-city Catholic schools. *Hispanic Journal of Behavioral Sciences, 15,* 310–323. doi:10.1177/07399863930153002

Van den Noortgate, W., & Onghena, P. (2003). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly, 18,* 325–346. doi:10.1521/scpq.18.3.325.22577

Varner, F., & Mandara, J. (2013). Differential parenting of African American adolescents as an explanation for gender disparities in achievement. *Journal of Research on Adolescence*. Advance online publication. doi:10.1111/jora.12063

*Véronneau, M.-H., & Dishion, T. J. (2011). Middle school friendships and academic achievement in early adolescence: A longitudinal analysis. *Journal of Early Adolescence, 31,* 99–124. doi:10.1177/0272431610384485

*Violato, C. (1990). A longitudinal comparative study of the academic achievement of education and other undergraduate students. *Canadian Journal of Education, 15,* 264–276. doi:10.2307/1495146

*von Wittich, B. (1972). The impact of the pass-fail system upon achievement of college students. *Journal of Higher Education, 43,* 499–508. doi:10.2307/1978896

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250–270. doi:10.1037/0033-2909.117.2.250

*Wang, J.-T., Tu, S.-Y., & Shieh, Y.-Y. (2007). A study on student performance in the college introductory statistics course. *AMATYC Review, 29,* 54–62.

*Wang-Cheng, R. M., Fulkerson, P. K., Barnas, G. P., & Lawrence, S. L. (1995). Effect of student and preceptor gender on clinical grades in an ambulatory care clerkship. *Academic Medicine, 70,* 324–326. doi:10.1097/00001888-199504000-00018

*Warren, J. S., Jackson, Y., & Sifers, S. K. (2009). Social support provisions as differential predictors of adaptive outcomes in young adolescents. *Journal of Community Psychology, 37,* 106–121. doi:10.1002/jcop.20273

*Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child Development, 62,* 1066–1078. doi:10.2307/1131152

*Whalen-Schmeller, B. (2006). Predicting English composition grades with the Academic Competence Evaluation Scales and the Multigroup Ethnic Identity Measure at an HBCU. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 67*(6), 2056.

*Whitley, J., Rawana, E. P., Pye, M., & Brownlee, K. (2010). Are strengths the solution? An exploration of the relationships among teacher-rated strengths, classroom behaviour, and academic achievement of young students. *McGill Journal of Education, 45,* 495–510. doi:10.7202/1003574ar

Wiborg, S. (2004). Education and social integration: A comparative study of the comprehensive school system in Scandinavia. *London Review of Education, 2,* 83–93. doi:10.1080/1474846042000229430

*Williams, T. R., Davis, L. E., Cribbs, J. M., Saunders, J., & Williams, J. H. (2002). Friends, family, and neighborhood: Understanding academic outcomes of African American youth. *Urban Education, 37,* 408–431. doi:10.1177/00485902037003006

Wilson, D. B. (2005). *Meta-analysis macros for SAS, SPSS, and Stata*. Retrieved from http://mason.gmu.edu/~dwilsonb/ma.html

*Witkow, M. R. (2009). Academic achievement and adolescents' daily time use in the social and academic domains. *Journal of Research on Adolescence, 19,* 151–172. doi:10.1111/j.1532-7795.2009.00587.x

*Witt, E. A., Dunbar, S. B., & Hoover, H. D. (1994). A multivariate perspective on sex differences in achievement and later performance among adolescents. *Applied Measurement in Education, 7,* 241–254. doi:10.1207/s15324818ame0703_6

*Woodfield, R., Jessop, D., & McMillan, L. (2006). Gender differences in undergraduate attendance rates. *Studies in Higher Education, 31,* 1–22. doi:10.1080/03075070500340127

*Woosley, S. A. (2005). Survey response and its relationship to educational outcomes among first-year college students. *Journal of College Student Retention: Research, Theory & Practice, 6,* 413–423.

*Wright, C. R., & Houck, J. W. (1995). Gender differences among self-assessments, teacher ratings, grades, and aptitude test scores for a sample of students attending rural secondary schools. *Educational and Psychological Measurement, 55,* 743–752. doi:10.1177/0013164495055005005

*Wright, R. E., & Palmer, J. C. (1998). Predicting performance of above and below average performers in graduate business schools: A split sample regression analysis. *Educational Research Quarterly, 22,* 72–79.

*Wright, R. J., & Bean, A. G. (1974). The influence of socioeconomic status on the predictability of college performance. *Journal of Educational Measurement, 11,* 277–284. doi:10.1111/j.1745-3984.1974.tb01000.x

*Wynne, W. D. (2003). An investigation of ethnic and gender intercept bias in the SAT's prediction of college freshman academic performance. *Dissertation Abstracts International: Section B. Sciences and Engineering, 64*(12), 2056.

*Yang, B., & Lu, D. R. (2001). Predicting academic performance in management education: An empirical investigation of MBA success. *Journal of Education for Business, 77,* 15–20. doi:10.1080/08832320109599665

*Young, I. P., & Young, K. H. (2010). An assessment of academic predictors for admission decisions: Do applicants varying in national origin and in sex have a level playing field for a doctoral program in educational leadership? *Educational Research Quarterly, 33,* 39–51.

*Young, J. W. (1991). Gender bias in predicting college academic performance: A new approach using item response theory. *Journal of Educational Measurement, 28,* 37–47. doi:10.1111/j.1745-3984.1991.tb00342.x

*Zarb, J. M. (1981). Non-academic predictors of successful academic achievement in a normal adolescent sample. *Adolescence, 16,* 891–900.

*Zeidner, M. (1987). A cross-cultural test of sex bias in the predictive validity of scholastic aptitude examinations: Some Israeli findings. *Evaluation and Program Planning, 10,* 289–295. doi:10.1016/0149-7189(87)90041-3

Received August 21, 2012
Revision received February 7, 2014
Accepted March 18, 2014 ■