

# Understanding the MBA Gender Gap: Women Respond to Gender Norms by Reducing Public Assertiveness but Not Private Effort

Personality and Social Psychology Bulletin  
 2017, Vol. 43(8) 1150–1170  
 © 2017 by the Society for Personality and Social Psychology, Inc  
 Reprints and permissions:  
[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)  
 DOI: 10.1177/0146167217708574  
[journals.sagepub.com/home/pspb](http://journals.sagepub.com/home/pspb)  


**Aaron S. Wallen<sup>1</sup>, Michael W. Morris<sup>1</sup>, Beth A. Devine<sup>2</sup>, and Jackson G. Lu<sup>1</sup>**

## Abstract

Women's underperformance in MBA programs has been the subject of recent debate and policy interventions, despite a lack of rigorous evidence documenting when and why it occurs. The current studies document a performance gap, specifying its contours and contributing factors. Two behaviors by female students that may factor into the gap are public *conformity* and private *internalization*. We predicted that women conform to the norm associating maleness with technical prowess by minimizing their public assertiveness in class discussions and meetings, but that they do not internalize the norm by reducing private effort. Data from multiple cohorts of a top-ranked MBA program reveal female underperformance occurred in technical subjects (e.g., accounting), but not social subjects (e.g., marketing). As predicted, the gender effect ran *not* through private effort but through public assertiveness, even controlling for gender differences in interests and aptitudes. These findings support some current policy interventions while casting doubt on others.

## Keywords

gender diversity, norms, performance, business education, assertiveness

Received March 12, 2016; revision accepted April 8, 2017

Only 15% of executives and 5% of CEOs are women, even though large studies find that firm performance increases with female executives (Jeong & Harrison, 2016; Nolan, Moran, & Kotschwar, 2016). Business schools may be to blame. Journalists report that at top MBA programs, women fare worse in grades and job offers (Kantor, 2013) and in early career trajectories (Kitroeff & Rodkin, 2015). However, past scientific research on MBA grades has found no gender differences (DeRue, 2009; Hancock, 1999; Wright & Bachrach, 2003; Yang & Lu, 2001). We present two studies that examine grade differences more closely and identify a circumscribed gender gap, in technical, but not social, classes.

Our analyses trace the gap to structural demographic features of MBA programs and to students' behavioral responses to these features. The population that enters top business schools varies by gender regarding technical interests and aptitudes. Not surprisingly, these preexisting characteristics affect grades, contributing to the gender gap. But more interestingly, students' behavior during business school—their responses to the salient gender differences—may factor into grades above and beyond this. Top MBA programs are majority male, have been even more so traditionally, and retain "locker-room or boot-camp" norms that masculinize

prowess in technical subjects like finance (see "As Academic Gender Gap Declines," 2011; Gellman, 2015; Kantor, 2013). The way female students respond to this norm may be a second-order process adding to the performance gap—one that, once identified, could be targeted by interventions.

But how female MBA students may respond to the gender norm is an open question. Social psychology has identified two qualitatively different ways that people may respond to their group being associated with lower performance: public *conformity*, behavior outwardly consistent with the association, or private *internalization*, behavior resulting from inner acceptance of the association. This distinction is fundamental in research on norms and behavior. For instance, in Asch's (1951) experiments, participants conformed to the norm publicly but not privately, whereas in Sherif's (1936) experiments with inherently ambiguous stimuli, participants adhered to the

<sup>1</sup>Columbia University, New York, NY, USA

<sup>2</sup>INSEAD, Fontainebleau, France

## Corresponding Author:

Aaron S. Wallen, Columbia University, 2970 Broadway, 502D, Lewisohn Hall, New York, NY 10027, USA.  
 Email: aw2328@columbia.edu

norm not just in public but in private as well. Female MBA students may merely *conform* to the association of masculinity with technical acumen in their public actions (i.e., participating less assertively in classroom and related study group discussions). Or they may *internalize* the expectation and disinvest their effort from the domain (i.e., reducing private effort, such as hours studying and in private tutoring sessions). In two large-scale studies with complementary methods, we find that public conformity, not private effort, is the behavioral mechanism that contributes to the MBA gender gap.

## Norms and Behavior

A central idea in many social sciences—anthropology, sociology, jurisprudence, political science, and so on—is that people’s actions are affected by surrounding norms, patterns of behavior associated with social categories. Although much recent psychology and public health research has focused on the effects of subjective norms (people’s assumptions or perceptions about social patterns), most disciplines study the effects of objective norms (the patterns themselves; see Morris, Hong, Chiu, & Liu, 2015). Classic social psychology studies of conformity examined objective norms rather than subjective norms: for instance, Sherif (1936) and Asch (1951) manipulated the pattern that objectively existed in a group; Newcomb (1943) tracked changes in newcomers to a college where liberalism was objectively prevalent; and modern classics (Cialdini, Kallgren, & Reno, 1991; Crandall, 1988) in this tradition similarly focus on effects of exposure to objective patterns in the actor’s social context.

In the case of gender and academic performance, objective norms can directly affect women’s performance. Women underperform on technical problems and in science, technology, engineering, and mathematics (STEM) courses when in majority male environments (Inzlicht & Ben-Zeev, 2000; Murphy, Steele, & Gross, 2007). In a majority male environment, the routines and artifacts of the culture tend to reflect male culture, and these social cues make women feel that they do not belong (Cheryan, Plaut, Davies, & Steele, 2009). When men are not only the majority but also are associated with technical prowess, the environment is doubly threatening. When women are put in an environment where the role models embody the stereotype of males as technically proficient and females as not so, this reduces women’s confidence about technical classes (Cheryan, Siy, Vichayapai, Drury, & Kim, 2011). So, objective gender differences in one’s community have an impact, whether they are explicitly perceived or just unconsciously registered.

## Research Setting: Graduate Schools of Business

Over the past century, schools of business have become institutionalized at the world’s major universities. Primarily, they offer a 2-year graduate degree, the MBA. MBA

programs have a standard curricular structure: a first year largely consisting of “core” or required classes—ranging from marketing to finance—and a second year consisting primarily of electives. The top programs seek applicants with 4 to 9 years of postcollege work experience in addition to strong undergraduate performance and aptitude test scores, making admissions highly competitive.

MBA programs mix students from many countries and industry backgrounds into large classes, which are characterized by animated discussions, debates, and role-plays of business dilemmas (Yeaple, 2012). Although the program we studied admits over 500 students per year, they are assigned to fixed subgroups: “clusters” of approximately 70 students who go through core classes together, and within this “learning teams” of five who complete assignments together. Unlike other graduate programs that are more solitary experiences, MBA programs are intensely interactive communities characterized by strong norms.

## Gender Differences in Preexisting Characteristics

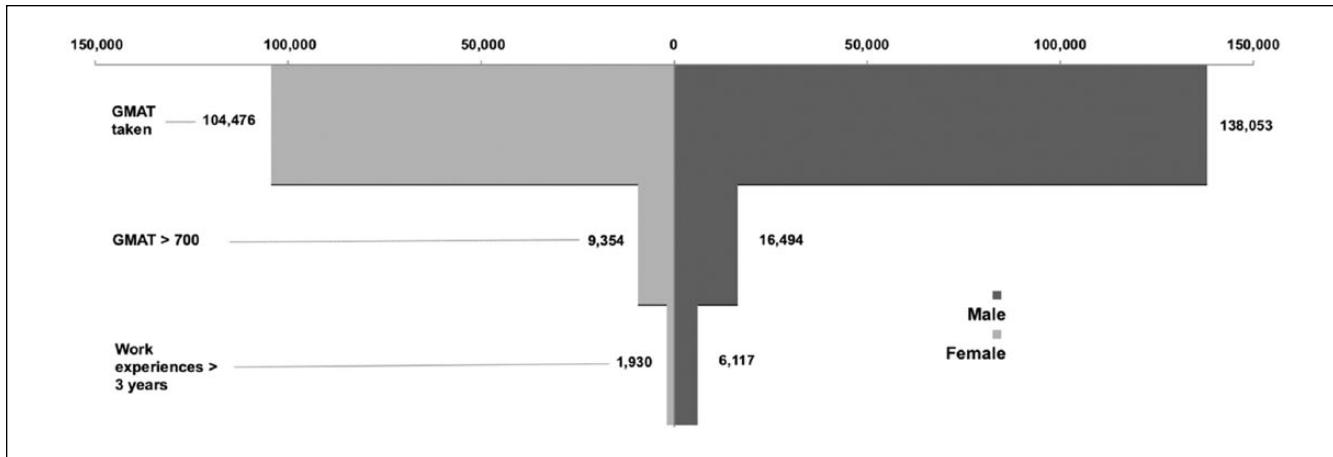
In considering the factors that may contribute to possible gender gaps in MBA student performance, the first category to consider are the characteristics with which students enter. Two relevant preexisting characteristics are interests and aptitudes, which are standard explanations for gender differences in higher education and career performance.

### Interests

Women are more likely to report interest in working with people than in working with numbers or things (Diekman, Brown, Johnston, & Clark, 2010; Else-Quest, Hyde, & Linn, 2010; Woodcock et al., 2013). Even among young children, females score higher on empathizing and males higher on systematizing (Baron-Cohen, 2003). Differences in interests are recognized in the MBA vernacular as the distinction between “poets,” interested in people and management, and “quants,” interested in numbers and finance (<http://poets-and-quants.com>). Baron-Cohen’s work, and other research on interests, suggests female MBA students may be less quant and more poet than their male peers. Because interests drive learning (Silvia, 2008), gender differences in quant interest could contribute to a gender gap in technical classes (Schmidt, 2011). Hence, the interest hypothesis is that differences in preexisting interests account for the gap in MBA student performance; women’s lower interest in technical topics contributes to their lower grades in related classes.

### Aptitudes

Applicants for MBA and other business-related graduate programs take the Graduate Management Admissions Test (GMAT), which assesses quantitative and verbal aptitudes.



**Figure 1.** The narrowing funnel of applicants to MBA programs, male and (especially) female, once filtered by high GMATs and high work experience.

Note. GMAT = Graduate Management Admissions Test.

The aptitude hypothesis, then, associated with Larry Summers among others, is that women are less frequently found at the highest levels of quantitative aptitudes, and this biological sex difference in quantitative aptitudes creates a gender gap in achievement in technical subjects.

But is the magnitude of sex differences in aptitudes sufficient to account for such a gap? Looking at the general population, sex differences in aptitudes are too slight to make an appreciable difference (Feingold, 1988; Hyde, Fennema, & Lamon, 1990; Hyde & Linn, 1988; Lindberg, Hyde, Petersen, & Linn, 2010). However, the population of students at elite business schools may be shaped by selection processes that differ by gender, creating gender differences in this subpopulation. Even if in the general population average GMATs are consistent across gender, if far fewer women than men with high GMATs enter the admissions pipeline for top-ranked MBA programs, these programs may end up enrolling classes that differ by gender in GMATs, despite using gender-blind admission standards. The key to this paradox is that the top programs would admit and compete for the same small pool of highly qualified female applicants. The result would be lower yields for highly qualified female than male admits, producing a gender difference in the enrolled class.

To explore whether this structural problem exists, we conducted a preliminary analysis on public archival data. We used the GMAT Interactive Profile tool available on the Graduate Management Admission Council (GMAC; n.d.) website. Access to the tool is granted to administrators and faculty of GMAC member business schools. We attained data from all GMAT examinees in the 2013-2014 testing year. The GMAT exam was taken 138,053 times by men and 104,476 times by women overall, indicating that many women pursue graduate programs related to business of one sort or another. The data indicate how the applicants plan to incorporate graduate education into their lives: full-time versus part-time programs and after how many years of prior work experience.

Examinees who fit the profile sought by top-ranked MBA programs, applying to full-time programs with 4 to 9 years of work experience, numbered 27,484 men to 13,141 women—more than twice as many men—a more male-skewed ratio than the overall ratio of examinees ( $Z = 40.62, p < .001$ ). By contrast, relatively more female (57%) than male GMAT-takers were applying to accounting graduate programs, which accept students right after college with no work experience (GMAC, 2015). The gender difference in choices about which programs to apply to, reflecting preferences about when and what to study, and ultimately, perhaps, societal structures such as work–family constraints, restricts the pipeline of women to top-ranked MBA programs.

But is the shortage sharp enough that it would create intense competition for the top-ranked female applicants? In other words, are there sufficient female applicants with the desired credentials to fill half the slots in the top-ranked MBA programs? The top programs seek students with GMAT scores over 700 and at least 4 years of work experience. Publicly available class profiles reveal that in all the top 15 programs, most students have these attributes. Of the 104,476 female students who took the GMAT in 2013-2014, 9,354 had GMATs greater than 700, but only 1,930 of them also had 4 or more years of work experience (Figure 1). Yet, to fill half the slots in the top 15 programs, approximately 7,000 women with these combined characteristics would be required. Even with the current policies of enrolling approximately 40 percent female students, the numbers suggest that the ratio of top-credentialed applicants to slots is much lower for females than males.

Our analysis of the applicant supply reveals a structure that would induce the elite programs to compete intensely for the highly qualified female applicants, and indeed, journalists have documented the competing offers of admission and scholarships that such female applicants receive (Gellman, 2015). The result would be lower enrollment yields for female

admits, and indeed many top-ranked schools report that they experience such a yield difference by gender (Columbia Business School, n.d.; Harvard Business School, n.d.-a, n.d.-b; Stanford Business School, 2015). The predictable result is that, despite gender-equal admissions, enrolled classes would differ by gender in GMATs, and so, the hypothesis that a gender gap results from differences in aptitudes is worth testing.

## Behavioral Responses to Gender Norm

In addition to their preexisting characteristics, students' behavior during the MBA program in response to the salient gender norm could independently contribute to a gender gap in performance. Two paths for this would be conformity—reduced public assertiveness—and internalization—reduced private effort.

### Public Assertiveness

Although measures of assertiveness generally show that women are lower than men, women's level of assertiveness depends a great deal on their situation. A meta-analysis found that American women's assertiveness rose and fell over the last century with changes in their social status (Twenge, 2001); for instance, when universities and workplaces became more female during World War II, women's level of assertiveness went up.

Studies suggest that women's adherence to gender norms about assertiveness in many settings reflect that these norms are enforced (highly assertive women are punished). In male-typed domains, such as negotiation, women attenuate their verbal assertiveness to avoid making negative social impressions (Amanatullah & Morris, 2010; Amanatullah & Tinsley, 2013). Women indicate they will speak up less than men when imagining working on a business team (Brescoll, 2012) or interviewing for business jobs (Moss-Racusin & Rudman, 2010).

In interviews, women at top MBA programs report fears of making negative social impressions by participating too actively in classes ("As Academic Gender Gap Declines," 2011). If women respond to the gender norm through *public* conformity, they would exhibit reduced verbal assertiveness in the classroom and study group meetings. Given MBA students' high concern about peer relationships, we expect that gender norms about technical prowess affect them in this way. Hence, our public conformity hypothesis is that less active participation by female students contributes to their lower grades. Less active participation would affect grades by reducing learning, as active participation is crucial to gaining competence, especially in technical subjects (Freeman et al., 2014).<sup>1</sup>

### Private Effort

Stereotype threat research proposes members of a group associated with lower performance sometimes internalize the message of lower proficiency and disinvest from the domain.

Steele's (1992) initial evidence for this mechanism of reduced private effort in academic underachievement was that SATs are less predictive of grades for Black than White college students. If female MBA students internalize negative performance expectations and withdraw their effort accordingly, their performance should similarly fall short of what is predicted by their GMATs. Studies measuring the process of reduced effort investment have found evidence that they contribute to minority underperformance in college grades (Owens & Massey, 2011). Although some studies find that women internalize particular gender norms as standards (Wood, Christensen, Hebl, & Rothgerber, 1997) or implicit attitudes (Nosek, Banaji, & Greenwald, 2002), studies do not find evidence for reduced effort as a contributor to gender differences in college grades (Cullen, Waters, & Sackett, 2006).

MBA students are known to be highly calculating in how they invest their time. This likely reflects their extensive training in economic models of rational choices. Studies find that economics training shifts decision making toward what is self-interest maximizing (see Frank, Gilovich, & Regan, 1993; Wang, Malhotra, & Murnighan, 2011). Reducing private effort would not be self-interest maximizing; it hurts learning and performance without boosting likeability, as it is largely unobservable to others. For these reasons, we do not expect to find support for the effort disinvestment hypothesis: that female MBAs lower their private effort.

Finally, a variant account in terms of lowered effort focuses on *single* female students, asserting that they allocate their time to dating the male students rather than academic effort (Patton, 2014). Although we fail to understand why such romance would take more time for the females than males involved, we also test if single female MBA students exert less academic effort than married female MBA students, or if female MBA students focus more on dating than male MBA students.

## Summary

Although differences in the interests and aptitudes that students bring to MBA programs could produce a gender gap in technical grades, the gap may grow larger through the ways students behaviorally respond to the association of gender with performance during the MBA program. Specifically, we propose that female students outwardly conform to the association of masculinity and technicality by reducing their public assertiveness in technical classes but that they do not internalize and disinvest by reducing their private effort. Therefore, our primary hypothesis across studies—the public conformity hypothesis—is that the gender difference in technical grades will be mediated by public assertiveness, even with alternative mediators included in the model, such as interests and aptitudes, and covariates, including ethnicity, citizenship status, financial aid, age, and undergraduate major.<sup>2</sup> These covariates reflect "third variables" that could be correlated with both gender and grades. For instance, students with a STEM or business undergraduate major have

more experience with technical courses and may be more likely to be male. The covariates were correlated—albeit weakly—with grades in our study, so it is important to control for these prior factors to isolate effects of our focal predictors.

To test our account, we conducted two studies with large, representative samples from two student cohorts enrolled in a top-ten MBA program. The two studies used datasets corresponding to different cohorts of students. The first study draws on archival data about students before and after enrollment in the program to identify the nature of the gender gap and directly test the public conformity hypothesis, both with and without simultaneous tests of the interests and aptitudes hypotheses. The second study uses a self-report survey to provide more fine-grained tests of the public conformity hypothesis, as well as a comparison with the disinvestment hypothesis.

## Study I

### Method

We received approval to compile and analyze all data for this study from the university's institutional review board. A dataset was created drawing student records from different offices of a business school, including admissions, registrar, student activities, and career services. These data were collected as part of the normal operation of the school, and they were analyzed as part of the school's efforts to monitor and improve its programs. We obtained grade data from the registrar's office. Whereas past studies have focused on the pre-existing characteristic of aptitude (GMATs), we added two measures of students' interests drawn from separate sources. Demographics, GMATs, and industry interests were drawn from admission records. Measures of broader life interests were drawn from a standard interest inventory involving hundreds of items that students complete for the career office's counseling program. In addition, we drew on students' assertiveness level scores that came from aggregated ratings by multiple classmates, conducted in a first-year class.

We created two key outcome variables—technical grade point average (GPA) and social GPA—by averaging grades from the core courses that loaded on these two factors. We then conducted an independent *t* test on each outcome variable to discover if academic outcomes varied by gender. We ran regression analyses to test which variables account for academic performance, controlling for other possible factors that might associate with gender. We conducted tests of indirect effects to determine which potential intervening variables helped explain the gender differences in grades.<sup>3</sup> To explore intervening mechanisms for gender differences, we computed indices of general life interests in “quant” versus “poet” topics as well as more specific industry interests. We also included quantitative and verbal GMAT scores, and an index of peer-rated assertiveness.

**Participants.** The graduating classes of 2009 and 2010 consisted of 1,324 students. Of these, we obtained complete data for 762 students. Some of the variables included in our model came from records that do not encompass every member of the student body, such as scores on the interest inventory used by our career counseling office, something in which not all students participate. To allow for accurate comparisons across statistical models, we only analyzed data for participants for whom we had full data in the most complete model. As a precaution, we conducted several analyses to rule out potential selection effects. First, although our sample of 762 participants had higher grades than the full sample of students, completion versus noncompletion of the interest inventory did not interact with gender to impact grades. Second, for each regression we report in the “Results” section, we relaxed the requirement for full data for all variables across the models and examined the results. The results based on the unrestricted dataset replicated those of the restricted dataset we have reported in the “Results” section. Therefore, we rule out some of the more potentially problematic impacts that selection may have had.

Given effect size conventions for multiple regressions (see Cohen, 1992), 762 participants allow us to detect even small effects. Furthermore, this sample size greatly exceeds samples in past MBA performance studies, typically in the low hundreds. Of the 762 participants, 496 were men and 266 were women.

### Covariates

**Ethnicity.** We created an indicator variable for each of the largest ethnic groups in the sample: African American or Black, East Asian, Latino, South Asian, and a category for other non-White groups. We coded these indicators as 0 if a student was not a member of the ethnic group and 1 if they were part of the group. Zeroes on all indicators indicated a White student.

**Citizenship status.** We treated citizenship as a dichotomous variable, with non-U.S. citizens coded as 0 and U.S. citizens coded as 1.

**Financial aid status.** We coded students who received some form of financial aid as 1, and students who did not as 0.

**Age.** We calculated age by subtracting students' dates-of-birth from the date on which students matriculated to school. The mean age for our sample was 27.52 ( $SD = 2.31$ , minimum = 21.67, maximum = 39.50).

**Undergraduate major.** We included two indicators of undergraduate major, STEM major and business major, each coded 1 if a student had the relevant major, and 0 if not.

### Predictor

**Gender.** We treated gender as a dichotomous variable, with men coded as 0 and women coded as 1.

### Potential intervening variables

**General interests.** Prior to MBA orientation, entering students are asked to complete an interest inventory used by career counselors (Butler & Waldroop, 2004). This inventory presents approximately 190 different tasks and work activities (e.g., coach a sports team, design a scientific experiment), and students indicate the level of enjoyment they anticipate for each item on a 4-point scale. Students receive feedback on an array of specific dimensions relevant to distinguishing business careers. Yet, conceptually, this broad range of items should reveal the two basic factors, corresponding to Baron-Cohen's (2003) systematizing and empathizing dimensions. Therefore, we conducted a principal components analysis (PCA) to formally assign Career Leader items to one of these two dimensions. We specified a solution with two principal components via a PCA with varimax rotation. The first component accounted for 12.9%, and the second component 10.9% of the total variance in the items after rotation.

We created composite scales using the highest loading items on each component, and each scale had high internal consistency: the first plainly a dimension of *quant* interests (e.g., "create a computer model for analyzing financial markets," "manage a portfolio of stocks for an investment company," "financial analyst," "use mathematical modeling to study weather systems," "mathematician," "create a complex airline route plan"; Cronbach's  $\alpha = .93$ ) and the second clearly a dimension of *poet* interests (e.g., "develop an advertising campaign for a product," "marketing brand manager," "lead a company task force studying personnel policy," "public relations professional," "counsel families in crisis," "train business professionals in team dynamics"; Cronbach's  $\alpha = .94$ ).

**Industry interests.** As part of their application for admissions, students had to indicate their desired industry interests. We collapsed their responses into three broad categories and created two indicator variables to capture these categories: Financial Services and Management. These correspond to the most common post-MBA industries, banking and management consulting. We coded the indicator as 0 if the student was not interested in a field and 1 if the student was interested in the field.

**GMATs.** As part of their applications for admission, students must submit scores from the GMAT. There are currently four sections of the exam, but we focus on the quantitative (GMATQ) and verbal (GMATV) sections only, as they are more established and reliable measures. Scores range from 0 to 60 on each section. In our sample, the mean GMATQ score was 46.50 ( $SD = 3.25$ , minimum = 33, maximum = 51), and the mean GMATV score was 40.64 ( $SD = 3.95$ , minimum = 25, maximum = 51).

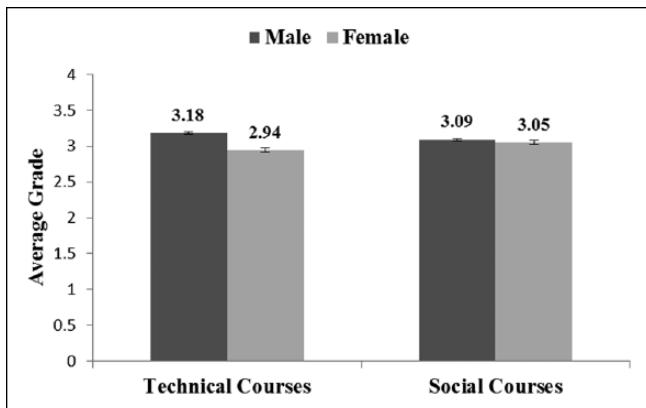
**Public assertiveness.** A central part of a first-year class on organizational behavior is rating classmates on their behav-

ioral and interpersonal styles. Peer ratings of this sort are familiar to students as they are regularly used in organizations for performance evaluation and development purposes. Students know that their ratings are kept anonymous, and only aggregated ratings from many classmates will be shared with the relevant student, so they take the task of giving feedback to their peers seriously, and give credence to how they are scored by their classmates. Raters are classmates from the same learning team and cluster who have worked on class assignments together. Each of the items was measured using a 7-point scale, with higher scores indicating agreement that the item's content was true of the target student. A composite scale to measure public assertiveness was created (Cronbach's  $\alpha = .85$ ). The scale items were "speaks up and shares own views when appropriate," "able to stand own ground in a heated conflict," "able to use vivid images and compelling logic and facts to support argument," and "willing to engage in constructive interpersonal confrontations."

### Outcomes

**Technical (GPA-T) and social course grades (GPA-S).** The school uses a 10-point scale when calculating GPAs, with 10 representing the highest grade (*H or Honors*), 7 the next highest (*HP or High Pass*), then 4 (*P or Pass*), 1 (*LP or Low Pass*), and 0 (*F or Failure*). However, for ease of understanding, we quantified grades using the standard 4-point scale: 4 for *H*, 3 for *HP*, 2 for *P*, 1 for *LP*, and 0 for *F*. Our goal was to ease interpretation of means and unstandardized regression coefficients. We multiplied numbers corresponding to grades by course credits, and then summed across courses, and then divided by the total credits to calculate GPA.

The dependent variable used in most studies of MBA academic performance—overall GPA—may hide critical differences, as grades in elective classes are negatively skewed, with most students receiving grades in the topmost categories. Some studies address this problem by calculating a GPA for the required "core" courses, which are standardized and strictly graded. However, even this more-valid measure of performance in competitive classes misses a key distinction in the eyes of MBAs and recruiters with regard to academic performance—performance in technical, quantitative courses versus social, people-oriented courses. Therefore, we distinguished between courses focused on topics where the material is complex technologically as opposed to complex psychologically. The students in our sample completed the following core, or required, courses: macroeconomics (t), microeconomics (t), accounting I (t), accounting II (t), statistics (t), decision modeling (t), corporate finance (t), operations (t), marketing I (s), corporate strategy (s), organizational change (s), organizational behavior (s), and marketing II. A confirmatory factor analysis established that a two-factor model,  $\chi^2 = 98.81$ ,  $df = 53$ , comparative fit index (CFI) = 0.98, Tucker-Lewis index (TLI) = 0.97, root mean square error of approximation (RMSEA) = 0.03, and standardized root mean square residual (SRMR) =



**Figure 2.** Average GPA-T and GPA-S for men and women (Study 1).

Note. The GPA-T effect is significant ( $p < .001$ ). The scale ranges from 0 to 4. Error bars represent SEs.  $n_{\text{male}} = 496$ ,  $n_{\text{female}} = 266$ . GPA-T = grade point average—technical; GPA-S = grade point average—social grades.

0.04—with courses marked with a “t” associated with the technical factor, and courses marked with an “s” associated with the social factor—fit the data better than a single-factor model,  $\chi^2 = 178.61$ ,  $df = 54$ , CFI = 0.94, TLI = 0.93, RMSEA = 0.06, and SRMR = 0.05; for the difference between these two nested models:  $\chi^2 = 79.80$ ,  $df = 1$ ,  $p < .001$ . Marketing II, which focuses on analytics, was conceptually associated with both factors, and thus was not included in the CFA. We averaged grades for the technical subject courses to create a GPA-T scale ( $M = 3.10$ ,  $SD = 0.49$ , minimum = 1.75, maximum = 4.00), and grades for social subject courses to create a GPA-S scale ( $M = 3.07$ ,  $SD = 0.46$ , minimum = 1.67, maximum = 4.00).

## Results

We report bivariate correlations in Table 1 and multiple regression results in Table 2.

**Gender.** As predicted, women ( $M = 2.94$ ,  $SD = 0.45$ ) received lower grades than men ( $M = 3.18$ ,  $SD = 0.48$ ) in technical courses, GPA-T,  $t(760) = -6.91$ ,  $p < .001$ , Cohen’s  $d = -.53$ , 95% confidence interval (CI) = [0.18, 0.32]. By contrast, women ( $M = 3.05$ ,  $SD = 0.44$ ) and men ( $M = 3.09$ ,  $SD = 0.44$ ) performed equally well in social courses, GPA-S,  $t(760) = -1.30$ ,  $p = .195$ ,  $d = -.10$ , 95% CI = [-0.02, 0.11] (Figure 2). Because a gender difference only existed on technical course grades, we conducted further analyses using this outcome only.

**Interests.** We tested whether interests were associated with technical course grades using two kinds of interest variables: general interests<sup>4</sup> and industry interests. Consistent with the interest hypothesis, there was an indirect effect of gender on technical course grades through general interests: overall effect = -.08,  $SE = .02$ , 95% CI = [-0.11, -0.05]; quant effect = -.06,  $SE = .01$ , 95% CI = [-0.09, -0.03]; and

poet effect = -.02,  $SE = .01$ , 95% CI = [-0.04, -0.01] (see Table 3, which shows the full statistical results for all indirect effects and their component paths). Also consistent with the interest hypothesis, there was an indirect effect of gender on technical course grades through industry interests: overall effect = -.02,  $SE = .01$ , 95% CI = [-0.05, 0.00]; Financial Services effect = -.02,  $SE = .01$ , 95% CI = [-0.05, -0.01]; and Management effect = .00,  $SE = .01$ , 95% CI = [-0.01, 0.02].

**GMATs.** Consistent with the aptitude hypothesis, there was a significant indirect effect of gender on technical performance through GMATs: overall effect = -.07,  $SE = .02$ , 95% CI = [-0.10, -0.04]; GMATQ effect = -.05,  $SE = .01$ , 95% CI = [-0.08, -0.03]; and GMATV effect = -.02,  $SE = .01$ , 95% CI = [-0.04, -0.00].

**Public assertiveness.** Consistent with the public conformity hypothesis, there was a significant indirect effect of gender on technical course performance through public assertiveness: indirect effect = -.03,  $SE = .01$ , bootstrapped 95% CI = [-0.06, -0.02]. To address our question of whether female students’ behavior during the MBA program incrementally contributes to the gender gap, we tested whether measures of behavioral responses predict GPA-T over and above measures of interests and aptitudes. Importantly, public assertiveness mediated the gender gap even with interests and GMATs included in the model as alternative mediators: indirect effect = -.03,  $SE = .01$ , bootstrapped 95% CI = [-0.06, -0.02] (Figure 3). This can be compared with the effects of the preexisting attributes that were entered simultaneously: quant interest, -.04,  $SE = .01$ , CI = [-0.06, -0.02]; poet interest, -.01,  $SE = .01$ , CI = [-0.02, 0.00]; Financial Services interest, -.02,  $SE = .01$ , CI = [-0.05, -0.01]; Management interest, .00,  $SE = .01$ , CI = [-0.01, 0.01]; GMATQ, -.05,  $SE = .01$ , CI = [-0.08, -0.03]; and GMATV, -.02,  $SE = .01$ , CI = [-0.03, -0.00].

## Discussion

Results show a circumscribed gender gap: female MBA students fare worse in technical courses but not in social courses. Lower public assertiveness during the MBA program contributed to this, even controlling for the contribution of pre-existing differences in GMATs and interests (including both general life interests and specific industry interests).

Some might question whether lowered assertiveness among female students during classes indicates their response to the program as opposed to a manifestation of their preexisting lower technical interests and GMATs. Could lower assertiveness among women just be a lingering reflection of preexisting low interest in all things technical? Could it be that lower aptitude female students feel they have less to contribute and that is what brings down the average female assertiveness? To test these alternative

**Table 1.** Correlation Table, Study 1.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Technical GPA																			
2. Social GPA	.43**																		
3. Black	-.26**	-.12**																	
4. East Asian	-.04	-.25**	-.10**																
5. Latino	.00	-.06	-.06	-.11**															
6. South Asian	-.02	.03	-.07	-.13**	-.07*														
7. Other	.06	.03	-.06	-.12**	-.07	-.08*													
8. U.S. citizen	-.18**	.07*	.12**	-.15**	-.14**	.05	-.17**												
9. Financial aid	-.17**	-.05	.22**	-.06	-.04	-.04	-.04	.24**											
10. Age	-.13**	-.04	.06	.04	.00	-.03	-.02	-.08*	-.07										
11. STEM major	.19**	.05	.04	-.05	.04	.11**	.09*	-.11**	-.03	.11**									
12. Business major	.07	-.06	-.05	.08*	.01	.01	-.03	-.03	-.02	-.15**	-.58**								
13. Female	-.24**	-.05	-.03	.14**	-.04	.04	-.03	.11**	.09*	-.12**	-.14**	-.03							
14. Poet	-.16**	.03	.04	.01	.01	.01	.00	.08*	.12**	.02	-.11**	-.08*	.20**						
15. Quant	.23**	-.01	.00	-.03	.06	-.01	.03	-.09**	-.05	.08*	.19**	-.02	-.32**	.21**					
16. Financial Services	.12**	-.03	.05	.02	.03	-.05	-.02	.03	-.07*	.03	-.10**	.21**	-.13**	-.18**	.25**				
17. Management	-.04	.00	-.05	.00	.00	.01	-.03	-.07	.02	.06	.06	-.06	-.01	.05	-.05	-.37**			
18. GMATQ	.40**	.01	-.18**	.24**	.02	.05	-.01	-.24**	-.16**	.02	.23**	.00	-.18**	-.14**	.20**	.07*	.01		
19. GMATV	.20**	.26**	-.18**	-.13**	-.15**	.01	.03	.17**	.04	-.02	-.08*	-.02	-.04	.04	.00	.03	.01	-.12**	
20. Assertiveness	.24**	.22**	-.04	-.26**	.03	-.04	.00	.03	-.05	.00	.03	-.04	-.21**	.06	.14**	.05	-.04	-.01	.16**

Note. Correlations within the ethnicity (Black, East Asian, Latino, South Asian, Other) and undergraduate major (STEM, Business) indicator variables are phi coefficients. Correlations among such indicator variables cannot be interpreted in a meaningful way, and are merely reflective of our sample characteristics (e.g., high “neither Black nor East Asian” frequencies, relatively low Black and East Asian frequencies, and—due to our coding such cases as “other”—zero “both Black and East Asian” frequencies). Correlations between dichotomous (e.g., Female) and continuous (e.g., Technical GPA) variables are point-biserial correlations. Correlations among continuous variables are Pearson product-moment correlations.  $N = 762$ . GPA = grade point average; GMATQ = GMAT-quantitative; GMATV = GMAT-verbal; GMAT = Graduate Management Admissions Test.

\* $p < .05$ . \*\* $p < .01$ .

**Table 2.** Multiple Regression Models of Technical Course Performance, Study 1.

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8		Model 9		
	b	SE	b	SE	b	SE	b	SE	b	SE									
African American/Black	-0.61***	0.08	-0.52***	0.08	-0.55***	0.08	-0.56***	0.07	-0.55***	0.07	-0.57***	0.07	-0.34***	0.07	-0.53***	0.07	-0.34***	0.07	
East Asian	-0.09	0.05	-0.13**	0.05	-0.14**	0.05	-0.10*	0.04	-0.10*	0.04	-0.10*	0.04	-0.15***	0.04	-0.03	0.05	-0.10*	0.04	
Latino	-0.05	0.07	-0.12	0.07	-0.15*	0.07	-0.14*	0.07	-0.15*	0.07	-0.15*	0.07	-0.07	0.06	-0.14*	0.07	-0.07	0.06	
South Asian	-0.09	0.06	-0.10	0.06	-0.18**	0.06	-0.15*	0.06	-0.14*	0.06	-0.15*	0.06	-0.17**	0.06	-0.13*	0.06	-0.13*	0.05	
Other	0.05	0.07	-0.02	0.07	-0.06	0.07	-0.05	0.06	-0.04	0.06	-0.05	0.06	-0.03	0.06	-0.03	0.06	-0.02	0.06	
U.S. citizen			-0.17***	0.04	-0.13**	0.04	-0.11**	0.04	-0.10**	0.04	-0.12**	0.04	-0.10**	0.04	-0.11**	0.04	-0.10**	0.04	
Financial aid			-0.12**	0.04	-0.11*	0.04	-0.10*	0.04	-0.09*	0.04	-0.09*	0.04	-0.09*	0.04	-0.09*	0.04	-0.06	0.04	
Age			-0.03***	0.01	-0.03***	0.01	-0.03***	0.01	-0.03***	0.01	-0.03***	0.01	-0.03***	0.01	-0.03***	0.01	-0.03***	0.01	
STEM Major					0.38***	0.05	0.33***	0.05	0.27***	0.05	0.33***	0.05	0.26***	0.04	0.33***	0.04	0.22***	0.04	
Business Major					0.23***	0.04	0.20***	0.04	0.16***	0.04	0.17***	0.04	0.17***	0.04	0.20***	0.04	0.15***	0.04	
Female							-0.19***	0.03	-0.13***	0.04	-0.18***	0.03	-0.13***	0.03	-0.16***	0.03	-0.06	0.03	
Poet									-0.09**	0.03							-0.07*	0.03	
Quant										0.15***	0.03							0.09**	0.03
Financial Services												0.10**	0.04					0.03	0.03
Management												-0.04	0.04					-0.04	0.04
GMATQ														0.05***	0.01			0.05***	0.01
GMATV														0.03***	0.00			0.03***	0.00
Assertiveness																0.13***	0.02	0.11***	0.02
R <sup>2</sup>	.08***		.13***		.20***		.24***		.26***		.25***		.36***		.27***		.39***		
ΔR <sup>2</sup>	.08***		.05***		.08***		.03***		.02***		.01**		.12***		.03***		.16***		

Note. Unstandardized regression coefficients (b) and SE are displayed in columns. The following predictors are coded as 1 if a student is of the labeled race/ethnicity, and 0 otherwise: African American/Black, East Asian, Latino, South Asian, Other. U.S. citizen is coded as 1 if a student is a U.S. citizen and 0 otherwise. Financial aid is coded as 1 if a student received financial aid, and 0 otherwise. Age is scaled in years. Female is coded as 1 if the student is a woman and 0 if a man. GMATQ, GMATV, Poet, Quant, and Assertiveness are treated as continuous predictors and are unit scaled. Financial Services and Management are coded as 1 if a student indicated interest in that field, and 0 otherwise. R<sup>2</sup> is the variance accounted for in technical course grades by the predictors in a given model. ΔR<sup>2</sup> is the change in variance accounted for due to the addition of predictors going from Model N – 1 to Model N for Models 1 through 4, and from Model 4 to Model N for Models 5 to 9. n<sub>male</sub> = 496, n<sub>female</sub> = 266. GMATQ = GMAT-quantitative; GMATV = GMAT-verbal; GMAT = Graduate Management Admissions Test.

\*p < .05. \*\*p < .01. \*\*\*p < .001.

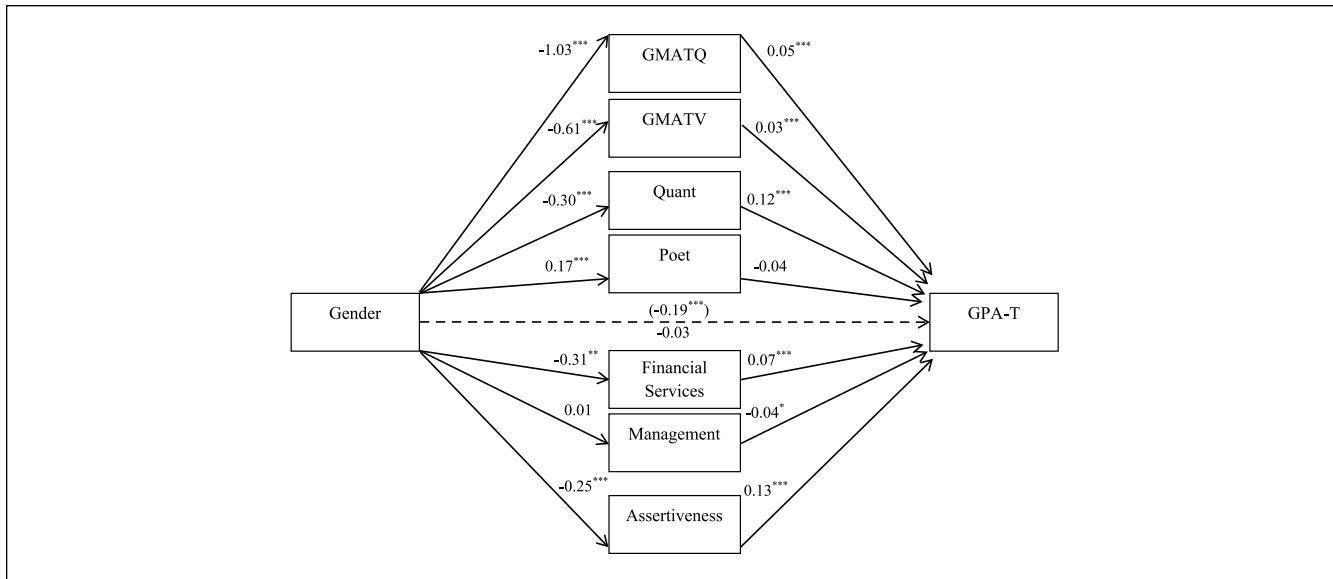
**Table 3.** Potential Intervening Variables of the Relationship Between Gender and Technical Course Performance, Study 1.

Mediator	Sex to mediator path					Mediator to GPA-T path					Indirect Effect			
	b <sup>a</sup>	SE	p	CIL <sup>a</sup>	CIU <sup>a</sup>	b	SE	p	CIL	CIU	Effect	SE	CIL	CIU
General interests														
Quant	-0.30	0.04	<.001	-0.38	-0.23	0.15	0.03	<.001	0.09	0.22	-0.06	0.01	-0.09	-0.04
Poet	0.17	0.04	<.001	0.09	0.24	-0.09	0.03	.008	-0.15	-0.02	-0.05	0.01	-0.07	-0.03
Industry interests														
Financial Services	-0.31	0.11	.006	-0.53	-0.10	0.07	0.02	<.001	0.03	0.11	-0.02	0.01	-0.05	-0.01
Management	0.01	0.13	>.250	-0.27	0.25	-0.04	0.02	.046	-0.08	0.00	0.00	0.01	-0.01	0.02
GMATs														
GMATQ	-1.21	0.25	<.001	-1.69	-0.73	0.06	0.00	<.001	0.05	0.07	-0.05	0.01	-0.08	-0.03
GMATV	-0.34	0.30	>.250	-0.93	0.25	0.03	0.00	<.001	0.02	0.04	-0.02	0.01	-0.04	-0.00
Public assertiveness	-0.25	0.05	<.001	-0.35	-0.15	0.13	0.02	<.001	0.09	0.18	-0.03	0.01	-0.06	-0.02
Public assertiveness <sup>b</sup>	-0.25	0.05	<.001	-0.35	-0.15	0.13	0.02	<.001	0.09	0.18	-0.03	0.01	-0.06	-0.02

Note. Unstandardized regression coefficients (b), SE, and 95% CIL and CIU limits are displayed in columns. Each grouping corresponds to the relationship between gender and the respective intervening variable (Sex to Mediator Path), the intervening variable to GPA-T (Mediator to GPA-T Path), or the indirect effect of gender on GPA-T through the intervening variable (Indirect Effect). GPA-T = grade point average-technical; CIL = confidence interval lower; CIU = confidence interval upper; GMAT = Graduate Management Admissions Test; GMATQ = GMAT-quantitative; GMATV = GMAT-verbal.

<sup>a</sup>Because Financial Services and Management are dichotomous variables, probit regression was used to estimate coefficients for the Gender to Financial Services and Gender to Management paths.

<sup>b</sup>This row contains results for Public Assertiveness with General Interests, Industry Interests, and GMATs also entered as mediators in the model.



**Figure 3.** Mediation of the gender to GPA-T relationship (Study 1).

Note. Values are unstandardized regression coefficients obtained from a bootstrapped test of indirect effects. The dashed line represents the mediated path. The value in parentheses above the dashed line reflects the total effect. The value below the dashed line represents the direct effect. The GPA-T variable consists of grade point average in technical subjects on a 4-point scale. Gender is coded such that male = 0, female = 1.  $n_{\text{male}} = 496$ ,  $n_{\text{female}} = 266$ . GPA-T = grade point average—technical.

\*\* $p < .01$ . \*\*\* $p < .001$ .

accounts, we regressed public assertiveness on our set of control variables as well as gender, general interests, and GMATs. Inconsistent with these accounts, neither quant interests,  $b = .06$ ,  $SE = .05$ ,  $t = 1.27$ ,  $p = .205$ , 95% CI =  $[-0.04, 0.16]$ , nor GMATQ predicted public assertiveness,  $b = .01$ ,  $SE = .01$ ,  $t = 0.684$ ,  $p > .250$ , 95% CI =  $[-0.01, 0.02]$ . So students' levels of public assertiveness during first-year MBA classes (and gender differences therein) do not simply reflect the intensity of technical interests and aptitudes with which they arrived; public assertiveness levels are independent of the preexisting differences.

## Study 2

Complementing the archival analysis in Study 1, Study 2 used an anonymous self-report survey of students. We included questions to measure public assertiveness more precisely and to measure private effort, for which we had no archival measures in Study 1. Study 2 involved students in a recent cohort (class of 2014) whereas the first study involved cohorts 4 and 5 years earlier.

An additional goal of Study 2 was to probe the correlates of success in the technical core courses, for example, business students' confidence going forward in their careers. To that end, Study 2 included a measure of academic confidence as a second-year MBA student, and we expected that it would be associated with the first-year performance on which our studies focus.

Our analytic strategy was the same as that used in Study 1. We conducted independent  $t$  tests to explore whether

gender gaps existed, multiple regressions to examine which variables associated with grades, and analyses of indirect effects to determine which potential intervening variables accounted for any detected gender gaps in grades.

## Method

**Participants.** We surveyed students from the 2014 graduating class during the second year of their MBA program. Of students in the class, 313 provided complete data for the variables included in our complete model. Of these, 206 students were men and 107 were women.

### Covariates

**Ethnicity.** We created an indicator variable for each of the largest ethnic groups in the sample: African American or Black, Asian, Latino, and a category for other non-White groups. We coded these indicators as 0 if a student was not a member of the ethnic group and 1 if they were part of the group. Zeroes on all indicators indicated a White student.

**Citizenship status.** We treated citizenship as a dichotomous variable, with non-U.S. citizens coded as 0 and U.S. citizens coded as 1.

**Age.** We asked students their age at the time of the survey. The mean age for our sample was 28.80 ( $SD = 1.92$ , minimum = 24, maximum = 36).

**Undergraduate major.** We included two indicators of undergraduate major, STEM major and business major, each

coded 1 if a student had the relevant major, and 0 if not.

#### *Predictor*

**Gender.** We treated gender as a dichotomous variable, with men coded as 0 and women coded as 1.

#### *Potential intervening variables*

**Industry interests.** We coded students' desired industry interests as in Study 1. Due to concerns regarding survey brevity, we were not able to include items regarding quant and poet interests in this study.

**GMATs.** We asked students to provide their GMAT scores. Scores can range from 0 to 60 on each section. In our sample, the mean GMATQ score was 46.84 ( $SD = 3.56$ , minimum = 30, maximum = 59), and the mean GMATV score was 42.15 ( $SD = 4.02$ , minimum = 32, maximum = 60).

**General assertiveness.** To construct a general assertiveness scale (Cronbach's  $\alpha = .78$ ) consistent with the one used in Study 1, we included three of the four items used in the Study 1 Assertiveness scale in our survey for Study 2. The items comprising the scale were "I speak up and share my views when appropriate," "I am able to stand my own ground in a heated conflict," and "I am willing to engage in constructive interpersonal confrontations."

**Quantitative assertiveness.** We measured assertiveness in quantitative academic contexts specifically by averaging two items to form a Quantitative Assertiveness scale (Cronbach's  $\alpha = .65$ ): "I feel more comfortable participating in quantitative classes than nonquantitative classes" and "I take the lead on quantitative problem sets and projects for my learning team."

By measuring self-reported assertiveness across these contexts, rather than peer-rated assertiveness, we eliminated possible peer biases in reacting to assertive behaviors that may distort the measure, for instance, that a behavior enacted by a women could be interpreted differently than the same behavior by a man (Biernat & Manis, 1994).

**Effort.** To measure academic effort, we averaged three survey items to create a composite scale (Cronbach's  $\alpha = .73$ ): "I believe it is important to excel in academics at [business school]," "I am willing to work hard to achieve academic success at [business school]," "When (if) I receive grades that are below my expectations, I (would) react by working harder on that course."

#### *Outcomes*

**GPA-T and GPA-S.** For Study 2, we asked students to indicate the grades they received in core courses. We converted these to a standard 4-point scale as follows: 4 for *H*, 3 for *HP*, 2 for *P*, 1 for *LP*, and 0 for *F*. We summed across classes, and then divided by the possible points to calculate GPA. As in

Study 1, we distinguished between courses heavy in technical, quantitative content and those that deal with social, psychological matters. At the time of this survey, the core comprised a slightly changed series of courses: global economic environment I (t), managerial economics (t), financial accounting (t), managerial statistics (t), business analytics and decision models (t), corporate finance (t), operations (t), marketing I (s), corporate strategy (s), organizational behavior (s), and marketing II. Courses marked with a "t" were associated with the technical factor, while courses marked with an "s" were associated with the social factor. Marketing II was again conceptually associated with both factors, and thus not included. Following the procedure used previously, we averaged grades for the technical subject courses to create a GPA-T scale ( $M = 3.32$ ,  $SD = 0.41$ , minimum = 2.29, maximum = 4.00), and for social subject courses to create a GPA-S scale ( $M = 3.29$ ,  $SD = 0.43$ , minimum = 2.00, maximum = 4.00).

**Romantic, rather than academic, priorities.** Two single-item measures of investment in romantic, rather than academic, agendas were included in the survey. These items were "It is likely that I will find a partner from within the [business school] community during my time at [the business school]" and "Dating is not one of my priorities while I'm at [business school]." These questions were only presented to single students.

**Second-year confidence.** To measure students' confidence at the time of the survey, we averaged responses from three items to form a composite measure (Cronbach's  $\alpha = .77$ ): "Today, I feel confident that I can handle the academics here," the reverse scored form of "relative to my classmates, I think I am performing worse academically," and "I feel confident in my ability to earn high grades in quantitative classes."

#### *Results*

We report correlations in Table 4 and multiple regression results in Table 5.

**Gender.** Replicating the results of Study 1, women ( $M = 3.17$ ,  $SD = 0.37$ ) received lower grades than men ( $M = 3.40$ ,  $SD = 0.41$ ) in technical courses,  $t(311) = -4.70$ ,  $p < .001$ , Cohen's  $d = .56$ , 95% CI =  $[-0.32, -0.13]$ , but not in social courses (women  $M = 3.28$ ,  $SD = 0.42$ ; men  $M = 3.30$ ,  $SD = 0.43$ ),  $t(311) = -0.31$ ,  $p > .250$ ,  $d = .04$ , 95% CI =  $[-0.12, 0.08]$  (Figure 4). As in Study 1, we confined further analyses to technical performance only.

**Interests.** Consistent with the interest hypothesis, students' industry interests mediated gender differences in technical class performance. There was an indirect effect of gender on technical course grades through industry interests: overall

**Table 4.** Correlation Table, Study 2.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Technical GPA																	
2. Social GPA	.42**																
3. Black	-.14*	-.12*															
4. Asian	.01	-.18**	-.12*														
5. Latino	-.10	-.05	-.06	-.20**													
6. Other	-.06	-.01	-.03	-.12*	.00												
7. U.S. citizen	-.16**	.09	.11	-.28**	-.17**	.02											
8. Age	-.24**	-.23**	.04	.14*	.11	-.01	-.13*										
9. STEM major	.12*	.01	.04	.16**	.10	-.03	-.32**	.18**									
10. Business major	.02	.00	.05	-.06	.00	.10	.06	-.10	-.44**								
11. Female	-.26**	-.02	.01	.04	-.07	.06	.13*	-.07	-.07	-.08							
12. Financial Services	.14*	-.15**	.07	.03	-.03	.01	.06	.06	-.13*	.16**	-.22**						
13. Management	-.21**	-.03	-.03	.03	.01	-.08	-.06	.00	.01	-.07	.12*	-.32**					
14. GMATQ	.30**	.05	-.06	.20**	-.05	-.06	-.17**	-.02	.22**	-.10	-.19**	.03	-.04				
15. GMATV	.27**	.21**	-.07	-.09	-.04	-.01	.23**	-.19**	-.08	-.02	-.11	.02	-.18**	-.03			
16. General assertiveness	.12*	.25**	.07	-.21**	-.04	.05	.01	-.19**	.00	-.07	-.13*	-.02	-.07	-.09	.07		
17. Quantitative assertiveness	.43**	-.06	-.01	.16**	.09	.00	-.25**	.04	.30**	.06	-.22**	.25**	-.14*	.28**	.04	-.03	
18. Effort	.20**	.21**	.08	-.10	-.03	.01	.09	.05	.09	-.04	.12*	.08	-.03	-.06	-.02	.13*	.14*

Note. Correlations within the ethnicity (Black, Asian, Latino, Other) and undergraduate major (STEM, Business) indicator variables are phi coefficients. Correlations among such indicator variables cannot be interpreted in a meaningful way, and are merely reflective of our sample characteristics (e.g., high “neither Black nor Asian” frequencies, relatively low Black and Asian frequencies, and—due to our coding such cases as “other”—zero “both Black and Asian” frequencies). Correlations between dichotomous (e.g., Female) and continuous (e.g., Technical GPA) variables are point-biserial correlations. Correlations among continuous variables are Pearson product-moment correlations.  $N = 313$ . GPA = grade point average; GMATQ = GMAT-quantitative; GMATV = GMAT-verbal; GMAT = Graduate Management Admissions Test.

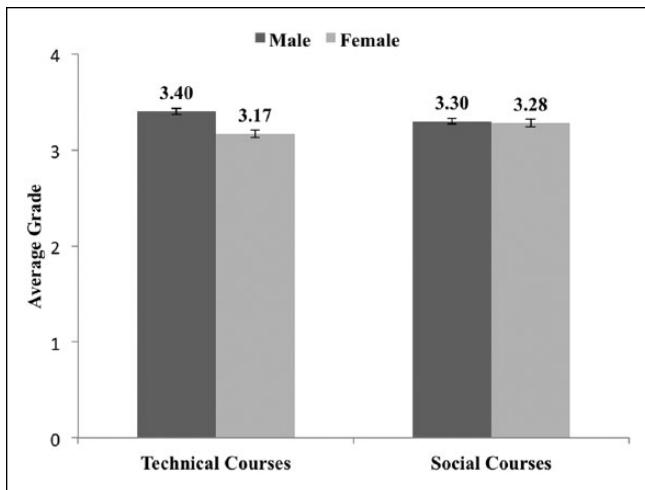
\* $p < .05$ . \*\* $p < .01$ .

**Table 5.** Multiple Regression Models of Technical Course Performance, Study 2.

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7		Model 8		Model 9		Model 10		
	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE	b	SE	
African American/ Black	-0.35**	0.13	-0.29*	0.12	-0.33**	0.12	-0.32**	0.12	-0.35**	0.12	-0.25*	0.11	-0.33**	0.12	-0.33**	0.11	-0.35**	0.11	-0.30**	0.10	
Asian	-0.03	0.05	-0.06	0.05	-0.07	0.05	-0.05	0.05	-0.05	0.05	-0.08	0.05	-0.04	0.05	-0.09	0.05	-0.02	0.05	-0.07	0.05	
Latino	-0.16*	0.08	-0.18*	0.08	-0.20*	0.08	-0.20*	0.08	-0.20**	0.08	-0.19*	0.07	-0.20*	0.08	-0.24**	0.07	-0.18*	0.08	-0.20**	0.07	
Other	-0.15	0.13	-0.15	0.13	-0.17	0.13	-0.12	0.12	-0.16	0.12	-0.11	0.12	-0.13	0.12	-0.13	0.11	-0.12	0.12	-0.14	0.10	
U.S. citizen			-0.18***	0.05	-0.14**	0.05	-0.11*	0.05	-0.13**	0.05	-0.15**	0.05	-0.11*	0.05	-0.08	0.05	-0.13**	0.05	-0.14**	0.04	
Age			-0.05***	0.01	-0.06***	0.01	-0.06***	0.01	-0.06***	0.01	-0.05***	0.01	-0.06***	0.01	-0.05***	0.01	-0.06***	0.01	-0.05***	0.01	
STEM major					0.17**	0.06	0.15**	0.06	0.16**	0.06	0.12*	0.05	0.15**	0.06	0.03	0.06	0.12*	0.06	0.01	0.05	
Business major						0.08	0.05	0.05	0.05	0.04	0.05	0.07	0.05	0.06	0.05	-0.02	0.05	0.05	0.05	0.01	0.04
Female							-0.21***	0.05	-0.18***	0.05	-0.14**	0.04	-0.21***	0.05	-0.15***	0.04	-0.24***	0.04	-0.12*	0.04	
Financial Services									0.08	0.05									0.00	0.04	
Management											-0.20**	0.06								-0.13*	0.05
GMATQ												0.03***	0.01							0.02***	0.01
GMATV												0.03***	0.01							0.02***	0.01
General assertiveness														0.03	0.03					0.02	0.03
Quantitative assertiveness																0.17***	0.02			0.12***	0.02
Effort																		0.14***	0.03	0.12***	0.03
R <sup>2</sup>	.04*		.13***		.15***		.21***		.25***		.31***		.21***		.34***		.27***		.45***		
ΔR <sup>2</sup>	.04*		.09***		.03*		.06***		.04***		.10***		.00		.13***		.06***		.25***		

Note. Unstandardized regression coefficients (b) and SE are displayed in columns. The following predictors are coded as 1 if a student is of the labeled race/ethnicity, and 0 otherwise: African American/Black, Asian, Latino, Other. U.S. citizen is coded as 1 if a student is a U.S. citizen and 0 otherwise. Age is scaled in years. Female is coded as 1 if the student is a woman and 0 if a man. GMATQ, GMATV, the three versions of Assertiveness, and Effort are treated as continuous predictors and are unit scaled. Financial Services and Management are coded as 1 if a student indicated interest in that field, and 0 otherwise. R<sup>2</sup> is the variance accounted for in technical course grades by the predictors in a given model. ΔR<sup>2</sup> is the change in variance accounted for due to the addition of predictors going from Model N - 1 to Model N for Models 1 through 4, and from Model 4 to Model N for Models 5 to 10. n<sub>male</sub> = 206, n<sub>female</sub> = 107. GMATQ = GMAT-quantitative; GMATV = GMAT-verbal; GMAT = Graduate Management Admissions Test.

\*p < .05. \*\*p < .01. \*\*\*p < .001.



**Figure 4.** Average GPA-T and GPA-S for men and women (Study 2).

Note. The GPA-T effect is significant ( $p < .001$ ). The scale ranges from 0 to 4. Error bars represent SEs.  $n_{\text{male}} = 206$ ,  $n_{\text{female}} = 107$ . GPA-T = grade point average—technical; GPA-S = grade point average—social grades.

effect =  $-.10$ ,  $SE = .04$ ,  $95\% \text{ CI} = [-0.20, -0.13]$ ; Financial Services effect =  $-.05$ ,  $SE = .03$ ,  $95\% \text{ CI} = [-0.12, -0.01]$ ; and Management effect =  $-.05$ ,  $SE = .03$ ,  $95\% \text{ CI} = [-0.13, -0.01]$  (see Table 6, which shows the full statistical results for all indirect effects and their component paths).

**GMATs.** As in Study 1, and consistent with the aptitude hypothesis, there was an indirect effect of gender on technical performance through GMATs: overall effect =  $-.07$ ,  $SE = .02$ ,  $95\% \text{ CI} = [-0.12, -0.03]$ ; GMATQ effect =  $-.04$ ,  $SE = .01$ ,  $95\% \text{ CI} = [-0.07, -0.01]$ ; and GMATV effect =  $-.03$ ,  $SE = .01$ ,  $95\% \text{ CI} = [-0.07, -0.01]$ .

**Assertiveness.** To test the public conformity hypothesis, we analyzed two different versions of assertiveness in Study 2: general assertiveness and quantitative assertiveness. There was no indirect effect of gender on technical performance through general assertiveness: indirect effect =  $-.01$ ,  $SE = .01$ , bootstrapped  $95\% \text{ CI} = [-0.03, 0.01]$ . However, there was an indirect effect through quantitative assertiveness: effect =  $-.07$ ,  $SE = .02$ , bootstrapped  $95\% \text{ CI} = [-0.12, -0.03]$ . These results, with respect to quantitative assertiveness, replicate findings from Study 1, and confirm our public conformity hypothesis that assertiveness would account for the gender gap in technical performance. To test whether the magnitude of assertiveness is substantial relative to other mediating factors, we tested whether it predicts performance with interests and GMATs included simultaneously as potential indirect paths. Even given this more inclusive model, there was an indirect effect through quantitative assertiveness: effect =  $-.06$ ,  $SE = .02$ , bootstrapped  $95\% \text{ CI} = [-0.10, -0.03]$  (Figure 5). This compares favorably with the effects of the preexisting attributes: Financial Services interest,

$-.05$ ,  $SE = .03$ ,  $95\% \text{ CI} = [-0.12, -0.01]$ ; Management interest,  $-.05$ ,  $SE = .03$ ,  $95\% \text{ CI} = [-0.13, -0.01]$ ; GMATQ,  $-.04$ ,  $SE = .01$ ,  $95\% \text{ CI} = [-0.07, -0.01]$ ; and GMATV,  $-.03$ ,  $SE = .01$ ,  $95\% \text{ CI} = [-0.07, -0.01]$ .

**Effort.** Reduced private effort by women, as would have been expected from the effort disinvestment hypothesis, did not contribute to the gender gap. In fact, results showed the opposite: women exerted slightly *more* effort than men,  $b = .19$ ,  $SE = .09$ ,  $p = .037$ ,  $95\% \text{ CI} = [0.01, 0.37]$ , and thus we do not find support for an explanation in which women reduced their effort after internalizing the norm. The analysis of indirect effects indicated an indirect effect through effort, effect =  $.03$ ,  $SE = .01$ , bootstrapped  $95\% \text{ CI} = [0.00, 0.06]$ , but this is a suppressor rather than a mediator of the gender gap. The effect of gender on technical course grades *increased* rather than decreased when effort was included in the model. Comparing Model 4 in Table 5,  $b = -.21$ ,  $SE = .05$ ,  $p < .001$ ,  $95\% \text{ CI} = [-0.30, -0.12]$ , with Model 9,  $b = -.24$ ,  $SE = .03$ ,  $p < .001$ ,  $95\% \text{ CI} = [-0.33, -0.16]$ , suggests effort suppresses the effect of gender on technical course grades. That is, if women had not exerted extra effort, the gender effect on grades would have been even larger. Again, based on this analysis, the data do not support an effort explanation for the gender gap in technical course grades.

Further evidence against female disinvestment in academic effort comes from the predictiveness of GMAT scores. A group that is disinvesting should have aptitude scores that are less predictive of performance (Steele, 1992). Instead, women's quantitative GMAT scores were as predictive of women's grades in technical courses,  $r(107) = .28$ ,  $p = .004$ , as men's quantitative GMAT scores were of men's performance in technical courses,  $r(206) = .27$ ,  $p < .001$ . This pattern is consistent with the effort scale findings; women are not failing to invest the effort needed to capitalize on their aptitude.

Nor did the data support the allegation that romantic agendas undermine academic performance for some female students. Single versus married status did not moderate the effect of gender on academic effort ( $p > .250$ ). Moreover, single women and men did not significantly differ in their responses to the item "It is likely that I will find a partner from within the business school community during my time at business school,"  $t(142) = 0.23$ ,  $p > .250$ , Cohen's  $d = .04$ ,  $95\% \text{ CI} = [-0.34, 0.43]$ . Similarly, single women and men did not differ on the item "Dating is not one of my priorities while at business school,"  $t(142) = 0.97$ ,  $p > .250$ , Cohen's  $d = .16$ ,  $95\% \text{ CI} = [-0.21, 0.62]$ .

**Technical performance and confidence.** As a final step, we tested whether there was a relationship between technical course grades and students' confidence as second-year students, including our set of control variables, gender, interests, GMATs, and quantitative assertiveness. Indeed, technical course grades significantly and positively predicted students' second-year confidence,  $b = .71$ ,  $SE = .10$ ,  $p < .001$ ,  $95\% \text{ CI}$

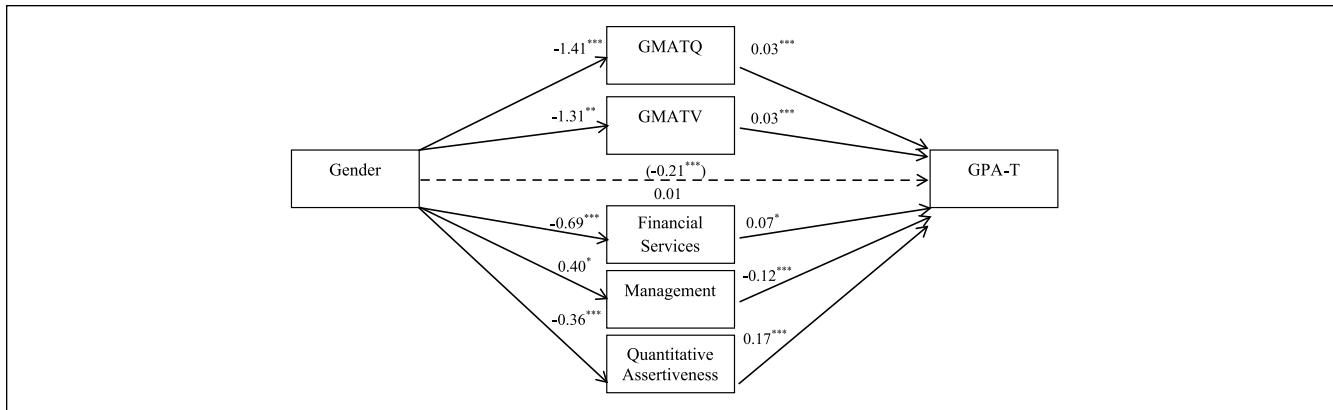
**Table 6.** Potential Intervening Variables of the Relationship Between Gender and Technical Course Performance, Study 2.

Mediator	Sex to mediator path					Mediator to GPA-T path					Indirect effect			
	b <sup>a</sup>	SE	p	CIL <sup>a</sup>	CIU <sup>a</sup>	b	SE	p	CIL	CIU	Effect	SE	CIL	CIU
Industry interests														
Financial Services	-0.69	0.18	<.001	-1.08	-0.36	0.07	0.03	<.001	0.01	0.12	-0.10	0.04	-0.20	-0.13
Management	0.40	0.20	.045	0.01	0.76	-0.12	0.03	.002	-0.19	-0.06	-0.05	0.03	-0.12	-0.01
GMATs														
GMATQ	-1.42	0.41	<.001	-2.23	-0.60	0.03	0.01	<.001	0.02	0.04	-0.07	0.02	-0.12	-0.03
GMATV	-1.26	0.46	.007	-2.17	-0.34	0.03	0.01	<.001	0.02	0.04	-0.03	0.01	-0.07	-0.01
General assertiveness	-0.19	0.08	.022	-0.35	-0.03	0.03	0.03	>.250	-0.04	0.09	-0.01	0.01	-0.03	0.01
Quant assertiveness	-0.41	0.11	<.001	-0.63	-0.19	0.17	0.02	<.001	0.13	0.21	-0.07	0.02	-0.12	-0.03
Quant assertiveness <sup>b</sup>	-0.36	0.11	<.001	-0.56	-0.17	0.17	0.02	<.001	0.13	0.21	-0.06	0.02	-0.10	-0.03
Effort	0.19	0.09	.037	0.01	0.37	0.15	0.03	<.001	0.10	0.20	0.03	0.01	0.00	0.06

Note. Unstandardized regression coefficients (b), SE, and 95% CIL and CIU limits are displayed in columns. Each grouping corresponds to the relationship between gender and the respective intervening variable (Sex to Mediator Path), the intervening variable to GPA-T (Mediator to GPA-T Path), or the indirect effect of gender on GPA-T through the intervening variable (Indirect Effect). GPA-T = grade point average-technical; CIL = confidence interval lower; CIU = confidence interval upper; GMAT = Graduate Management Admissions Test; GMATQ = GMAT-quantitative; GMATV = GMAT-verbal.

<sup>a</sup>Because Financial Services and Management are dichotomous variables, probit regression was used to estimate coefficients for the Gender to Financial Services and Gender to Management paths.

<sup>b</sup>This row contains the results for Quant Assertiveness with Industry Interests and GMATs also entered as mediators in the model.



**Figure 5.** Mediation of the gender to GPA-T relationship (Study 2).

Note. Values are unstandardized regression coefficients obtained from a bootstrapped test of indirect effects. The dashed line represents the mediated path. The value in parentheses above the dashed line reflects the total effect. The value below the dashed line represents the direct effect. The GPA-T variable consists of grade point average in technical subjects on a 4-point scale. Gender is coded such that male = 0, female = 1.  $n_{\text{male}} = 206$ ,  $n_{\text{female}} = 107$ . GPA-T = grade point average—technical.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$=[0.52, 0.90]$ , suggesting that performance in core technical classes is associated with academic confidence a year later.

## Discussion

As in Study 1, women received lower grades than men in technical courses, but not in social courses. Again, this was mediated not only by differences in preexisting characteristics (interests and GMATs), but also over and above these by public assertiveness. Not surprisingly, assertiveness behavior specifically in technical courses mattered more than general assertiveness. Unlike the Study 1 version of assertiveness, which was other-rated and referenced shared classes, the self-rated version relies on a student's view of his or her own assertiveness. As these self-views are informed by many contexts, academic and nonacademic, the self-rated assessment in quantitative classes more closely captures our construct. Results did not show that women disinvest from technical courses as their GMAT levels do not overpredict performance relative to men. Indeed, self-reported academic effort was slightly higher among women than men. That said, we note that private effort was measured at a more general level than the outcome variable, which could reduce our ability to detect a significant negative indirect effect through effort. But given that we observed a significant *positive* indirect effect, it seems unlikely that the generality of the measure hindered the detection of a relationship.

To check that our effort findings are not just artifacts of gender differences in self-report biases, we added a follow-up—Study 2<sub>supp</sub>—investigating time-investment in free private tutoring sessions, by accessing records at the same MBA program. The tutoring sessions are almost exclusively focused on technical subjects. We tallied tutoring hours for first-year MBAs in the fall of 2014, and found that female students ( $M = 4.88$  hr) invested nearly 3 times as much as

men ( $M = 1.73$  hr) in studying with tutors,  $t(542) = 7.72$ ,  $p < .001$ , Cohen's  $d = .66$ , 95% CI = [2.23, 4.07]. So this objective measure of private effort corroborates the self-report measure effect that women put in more private effort.

## General Discussion

Results from two studies documented a gender gap in performance at a top MBA program in technically oriented classes but not socially oriented classes. Although focusing on cohorts 5 years apart in time, the gender gap was remarkably stable, in both cases a substantial difference in technical courses (roughly a quarter of a point on the standard 4-point GPA scale) and no difference in social courses. Study 1 found that this effect in technical classes was mediated by students' public assertiveness (as rated by their peers), even when controlling for a comprehensive set of preexisting differences. Study 2 replicated this finding with self-report measures and established it was assertiveness in technical classes, not assertiveness generally, that mattered—consistent with conformity to the norm associating maleness with technical prowess. Study 2 additionally showed that performance in technical core courses strongly predicted confidence as a second-year student, indicating its lasting consequences.

Study 2 also tested accounts centered on private effort. Inconsistent with such accounts, GMATs do not overpredict grades for women, which would have been the case if women disinvested effort from academics. Self-reported private effort was higher for women than men, and as such, private effort is a suppressor rather than a mediator of the gender gap. Study 2<sub>supp</sub> corroborates this with an objective measure of private effort—women spend 3 times as many hours in private tutoring sessions (perhaps asking all the questions that they have held back in public). Finally, contrary to the tired allegation that female students pursue a "Mrs. Degree," being

single did not affect women differently from men academically or in terms of energy invested into their social lives.

Although the past literature on MBA academic performance showed no evidence for a gender difference, the current studies provide a more rigorous test of the question attracting much recent debate: Is there a gender gap at the top MBA programs that feed the upper echelons of business? To our knowledge, the current research comprises the only published data about students' performance at a top-10 MBA program in many decades. They go beyond prior studies in measuring preexisting attributes potentially relevant to gender, such as quant versus poet interests. Also, past studies looked only at an overall measure of GPA thereby potentially missing a gap that arises solely in technical classes.

Several evidential advantages come from the combination of our two studies. First, although both studies sampled a high fraction of the entire student population, the students left out, and any associated biases, differ between archival and survey methods. For example, if highly conscientious students completed the Study 2 survey more often than less conscientious students, the sample would contain a disproportionate number of highly conscientious students compared with the school's MBA population. However, this self-selection bias would not affect the sample gathered from archival sources in Study 1. Second, these methods allowed us to operationalize key outcome variables in complementary ways: whereas Study 1 measures outwardly visible behavioral manifestations of assertiveness through peer ratings, Study 2 measures it through self-reports. Third, we replicated our findings with cohorts of students separated by a significant amount of time (5 years apart), finding a consistent pattern.

Study 1 is particularly strong compared with prior studies in the literature given that it entirely avoids self-report data and draws instead on disparate and high-validity archival measures. In study 1, we eliminate common method bias by assembling data from different sources: applications from the admissions office, a vocational inventory from the career counseling office, grades from the registrar, and peer ratings of assertiveness from a class assignment. Archival data also avoid the possible problem of selection bias by gender in prior studies: different subsets of male versus female students may agree to participate in studies. Archival data also avoid the social desirability and experimenter demand biases of self-report surveys. The archival measures we used are highly reliable: both GMATs and the vocational inventory are tests with hundreds of extensively validated items that students take very seriously; measures of academic performance and assertiveness aggregate inputs from multiple evaluators (professors and classmates, respectively), each of whom work from extensive observations rather than brief impressions. One difference between results in the two studies is the relative strength of each mediating path. This is indicated by the size of the indirect effects when the multiple mediators are entered simultaneously. In Study 1, assertiveness remained significant even when interest and aptitude measures were entered, but it was

not the strongest path. In Study 2, assertiveness was the strongest path. One possible explanation is that self-reports capture assertiveness levels better than peer ratings. For instance, peer ratings may register given level of class participation as more assertive when performed by a woman than when performed by a man. However, it is also the case that Study 1 had more and better measures of interests than Study 2, so another possibility is that the less dominant strength of assertiveness in Study 1 reflects this greater competition. In any case, the important finding in both studies is that an appreciable part of the gender gap in performance runs through the way students behaviorally respond to the MBA classroom, not just through preexisting gender differences in interests and aptitudes that they bring to MBA programs.

### ***Theoretical Implications***

These results indicate a gender gap in MBA performance that arises from several contributing factors—some that precede the MBA program and one that occurs in students' behavior during the program. Our preliminary analysis of GMAT examinees showed novel and important evidence for differential selection into the admissions pipeline to top MBA programs. This explains the gender difference in schools' competition for admittees and their enrollment yields. We suggest that this elucidates why gender-equal admissions can nonetheless result in gender-unequal enrolled classes in terms of GMATs. We consistently observed such gender differences in GMATs. These preexisting differences in students' technical aptitudes—as well as technical interests—did contribute to the gender gap.

But so did students' behavior during the MBA program. Consistent with a response of public conformity to the norm associating maleness with technical prowess, women showed less assertiveness, particularly in technical classes, and this contributed to the gender gap in grades. Women did not show any less private effort; in fact, they showed more private effort.

These findings of multiple determinants weigh against the simplistic, one-factor explanations that have figured in debates about gender differences in graduate school performance. Partisans who would reduce gender gaps to a single favored explanation—biological sex differences in aptitudes, societal structures that burden women with child care, or the sexist prejudice of professors—could not account for the multiple paths of influence that we observe.

### ***Policy Implications***

The last 2 years have seen many policy debates and interventions at business schools aimed at reducing gender gaps, both at the school we studied and many others. A recent conference of business school deans at the White House identified a set of best practices, including targeted outreach toward high school and undergraduate women to draw them into the MBA pipeline (Korn, 2014; The White House, Office of the

Press Secretary, 2015). One top-10 school has attempted to attract women applicants by featuring female students and prominent alumnae at events (Blackman, 2011). Many schools have announced new scholarships to lure talented women onto the MBA track (Gellman, 2015).

In addition to these pipeline-related fixes, other interventions have sought to change what happens during MBA programs. One example is another top-10 school's reported policy of encouraging faculty to assign grades in each given class that do not differ by gender (Kantor, 2013). The current findings that gender differences in grades in part reflect aptitudes and interests suggest pressure for grade equality is problematic. Though removing external signs of the gender norm could make the norm less salient, doing so distorts the signal that grades provide about performance and thereby risks cynicism and disillusionment. Signs of this can be gleaned from a student comment: "Diversity at [X] Business School means playing around with grades . . . rather than tackling the underlying issues" (Kitroeff, 2014). Solving the gender gap in performance by encouraging gender-equal grading is like responding to global warming by adjusting thermometer scales.

Our results suggest that there is nothing inherent about the gender gap and that the right interventions could reduce it. A successful intervention should address both the source of the gender norm and behavioral responses to it. It arises from objective gender differences in the student body, which reflect the structural problems, policies, and biased pipeline that our analysis of GMAT test takers has documented. An admissions policy change—enrolling 50% female classes—would help in taking away women's minority status. But this change alone will not solve the problem (and could even worsen it) unless it is paired with interventions to improve the gender ratio of the admissions pipeline. With the status quo applicant pool, our analysis suggests that heightened demand for high-credentialed female applicants would increase admissions competition for them, decrease the yield that programs enroll of such admittees, and, as an unintended byproduct, increase gender differences in the student body on dimensions such as GMATs and work experience. Although business schools cannot control all the societal factors that contribute to the pipeline problem, they would be wise to proliferate and publicize scholarships for women. They could sponsor more programs encouraging STEM courses for undergraduate women. Countries where schoolgirls score relatively highly in math tend to have more women in management positions (Nolan et al., 2016). A different approach is reducing admissions targets for prior work experience so students can enroll at a younger age—to become more like law schools which have more even gender ratios (Law School Admission Council, n.d.).

It is also important to address the mechanism through which the gender norm impacts behavior: public conformity through reduced assertiveness in technical classes. Interventions for assertiveness often focus on women's confidence or body language. But this misses the point. Public

conformity is not just a behavioral quirk to be trained out of people; it is a pragmatic response to a norm in the social environment (Amanatullah & Morris, 2010). And this norm creates serious consequences for perceived norm violations, as women who do not conform to gender norms risk backlash (Rudman, 1998) and interpersonal penalties (Heilman, Wallen, Fuchs, & Tamkins, 2004). Broader changes to organizational procedures can reduce gender norm enforcement (Sandberg, 2013). In MBA programs, this could mean training and evaluating male and female students for skills in eliciting participation from their female peers. Likewise, professors could be taught to better manage participation—the outcome of equal participation in a diverse group sometimes requires subtly unequal treatment of different subgroups by the facilitator. In sum, "lean in" interventions that address the roots of the problem would work better than "lean on" interventions that suppress its symptoms. Fortunately, several interventions at this stage are feasible, and we hope that, after implementing some of our recommendations, the flow of top women MBAs to corporate leadership soon resembles a deluge rather than a trickle.

### Authors' Note

This research was done as part of a Columbia Business School Leadership Lab Research Initiative on Antecedents of MBA Success. The second study was conducted as part of a gender initiative by a team of students from the class of 2014 overseen by Dean Katherine Phillips and advised regarding the survey design by Michael Morris.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and /or publication of this article: This work was supported in part by a Columbia Business School Dean's Office Interdisciplinary Research Grant.

### Supplemental Material

The supplemental material is available online.

### Notes

1. Another possibility is that level of public participation affects grades through professors' evaluation of class participation. However, given that the weight of class participation is low for technical courses (33% in social classes but only 13% in technical classes), this alternative path is unlikely.
2. Although specific coefficient values differed with the inclusion of covariates, in no case did the substantive results change.
3. To conduct tests of indirect effects involving continuous mediators, we used the Process Procedure for SPSS with 5,000 bootstrapped samples and percentile 95% confidence intervals (Hayes, 2013). We used M-Plus as a test of mediation for

- analyses that included categorical mediators and for the confirmatory factor analyses involving grades (Muthén & Muthén, 2012). Tests of indirect effects were conducted with each potential mediating variable, one at a time, to test how well each factor accounts for the gender difference. In addition, we tested the mediating role of our key variable, assertiveness, when entered simultaneously with all the other potential mediators (Graduate Management Admissions Test [GMATQ and GMATV], interests, etc.) to specify the incremental role of this behavioral factor in mediating the gender difference in performance.
4. For grouped intervening variables, for example, Quant & Poet, effects on grade point average-technical (GPA-T) control for the other intervening variable in the group. This is to keep the path coefficients reported in Tables 3 and 6 consistent with the associated analysis of indirect effects, wherein grouped variables were entered simultaneously as potential intervening variables.

## References

- Amanatullah, E. T., & Morris, M. W. (2010). Negotiating gender roles: Gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others. *Journal of Personality and Social Psychology*, 98, 256-267.
- Amanatullah, E. T., & Tinsley, C. H. (2013). Punishing female negotiators for asserting too much...or not enough: Exploring why advocacy moderates backlash against assertive female negotiators. *Organizational Behavior and Human Decision Processes*, 120, 110-122.
- As academic gender gap declines, there is still work to be done. (2011, April 25). *The Harbus*. Retrieved from <http://www.harbus.org/2011/gender-gap/>
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177-190). Pittsburgh, PA: Carnegie Press.
- Baron-Cohen, S. (2003). *The essential difference: Men, women and the extreme male brain*. London, England: Penguin.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66, 5-20.
- Blackman, S. (2011, August 19). More women head to school for M.B.A.s. *U. S. News & World Report*. Retrieved from <http://www.usnews.com/education/blogs/mba-admissions-strictly-business/2011/08/19/more-women-head-to-school-for-mbas>
- Brescoll, V. L. (2012). Who takes the floor and why: Gender, power, and volubility in organizations. *Administrative Science Quarterly*, 56, 622-641.
- Butler, T., & Waldrop, J. (2004). A function-centered model of interest assessment for business careers. *Journal of Career Assessment*, 12, 270-284.
- Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, 97, 1045-1060.
- Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science*, 2, 656-664.
- Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). The focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 201-234). San Diego, CA: Academic Press.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Columbia Business School. (n.d.). *Class profile*. Retrieved from <https://www8.gsb.columbia.edu/programs-admissions/mba/admissions/class-profile>
- Crandall, C. (1988). Social contagion of binge eating. *Journal of Personality and Social Psychology*, 55, 588-598.
- Cullen, M. J., Waters, S. D., & Sackett, P. R. (2006). Testing stereotype threat theory predictions for math-identified and non-math-identified students by gender. *Human Performance*, 19, 421-440.
- DeRue, D. S. (2009). Quantity or quality? Work experience as a predictor of MBA student success (GMAC Research reports RR-09-9). McLean, VA: GMAC.
- Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, 21, 1051-1057.
- 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.
- Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist*, 43, 95-103.
- Frank, R. H., Gilovich, T., & Regan, D. T. (1993). Does studying economics inhibit cooperation? *The Journal of Economic Perspectives*, 7, 159-171.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., ... Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111, 8410-8415.
- Gellman, L. (2015, May 6). Business schools are fighting to recruit top women. *The Wall Street Journal*. Retrieved from <http://www.wsj.com/articles/why-business-schools-are-fighting-over-top-women-1430957422>
- Graduate Management Admission Council. (2015). *Application Trends Survey* [Data file]. Available from <http://www.gmac.com/~/media/Files/gmac/Research/admissions-and-application-trends/2015-application-trends-web-release-1.pdf>
- Graduate Management Admission Council. (n.d.). *GMAT® Interactive Profile* [Data file] Available from <http://www.gmac.com/market-intelligence-and-research/gmat-statistics/interactive-profile-of-gmat-candidates.aspx>
- Hancock, T. (1999). The gender difference: Validity of standardized admission tests in predicting MBA performance. *Journal of Education for Business*, 75, 91-93.
- Harvard Business School. (n.d.-a). *Admissions class profile*. Retrieved from <http://www.hbs.edu/mba/admissions/class-profile/Pages/default.aspx>
- Harvard Business School. (n.d.-b). *Admissions policies*. Retrieved from <http://www.hbs.edu/mba/admissions/application-process/Pages/admission-policies.aspx>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.
- Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed

- at male gender-typed tasks. *Journal of Applied Psychology*, 89, 416-427.
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139-155.
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, 104, 53-69.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11, 365-371.
- Jeong, S. H., & Harrison, D. (2016). Glass breaking, strategy making, and value creating: meta-analytic outcomes of females as CEOs and TMT members. *Academy of Management Journal*. Advance online publication. doi:10.5465/amj.2014.0716
- Kantor, J. (2013, September 7). Harvard Business School case study: Gender equity. *The New York Times*. Retrieved from <http://www.nytimes.com/2013/09/08/education/harvard-case-study-gender-equity.html>
- Kitroeff, N. (2014, November 11). Best business schools 2014: Duke tops full-time MBA list, Harvard slips. *Bloomberg Businessweek*. Retrieved from <http://www.businessweek.com/articles/2014-11-11/best-business-schools-2014-duke-is-top-full-time-mba-harvard-slips>
- Kitroeff, N., & Rodkin, J. (2015). The real payoff from an MBA is different for men and women. *Bloomberg Businessweek*. Retrieved from <http://www.bloomberg.com/news/articles/2015-10-20/the-real-cost-of-an-mba-is-different-for-men-and-women>
- Korn, M. (2014, April 17). B-School deans get to work at the White House. *The Wall Street Journal*. Retrieved from <http://blogs.wsj.com/atwork/2014/04/17/b-school-deans-get-to-work-at-the-white-house/>
- Law School Admission Council. (n.d.). *Matriculants by ethnic and gender group* [Data file]. Retrieved from <http://www.lsac.org/lsacresources/data/ethnic-gender-matriculants>
- 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.
- Morris, M. W., Hong, Y. Y., Chiu, C. Y., & Liu, Z. (2015). Normology: Integrating insights about social norms to understand cultural dynamics. *Organizational Behavior and Human Decision Processes*, 129, 1-13.
- Moss-Racusin, C. A., & Rudman, L. A. (2010). Disruptions in women's self-promotion: The backlash avoidance model. *Psychology of Women Quarterly*, 34, 186-202.
- Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science*, 18, 879-885.
- Muthén, L. K., & Muthén, B. O. (2012). *Statistical analysis with latent variables: User's guide* (Version, 7). Retrieved from [http://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver\\_7\\_r3\\_web.pdf](http://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r3_web.pdf)
- Newcomb, T. M. (1943). *Personality and social change*. New York, NY: Holt, Rhinehart & Winston.
- Nolan, M., Moran, T., & Kotschwar, B. (2016, February). *Is gender diversity profitable? Evidence from a global survey* (Working Paper No. 16-3). Retrieved from <http://www.piie.com/publications/wp/wp16-3.pdf>
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, 83, 44-59.
- Owens, J., & Massey, D. S. (2011). Stereotype threat and college academic performance: A latent variables approach. *Social Science Research*, 40, 150-166.
- Patton, S. (2014). *Marry by choice, not by chance: Advice for finding the right one at the right time*. New York, NY: Gallery Books.
- Rudman, L. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, 74, 629-645.
- Sandberg, S. L. (2013). *Women, work, and the will to lead*. New York, NY: Alfred A Knopf.
- Schmidt, F. L. (2011). A theory of sex differences in technical aptitude and some supporting evidence. *Perspectives on Psychological Science*, 6, 560-573.
- Sherif, M. (1936). *The psychology of social norms*. New York, NY: Harper.
- Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science*, 17, 57-60.
- Stanford Business School. (2015). *Entering class profile*. Retrieved from <https://www.gsb.stanford.edu/programs/mba/admission/evaluation-criteria/class-profile>
- Steele, C. M. (1992, April 1). Race and the schooling of Black Americans. *The Atlantic*. Retrieved from <http://www.theatlantic.com/magazine/archive/1992/04/race-and-the-schooling-of-black-americans/306073/>
- Twenge, J. M. (2001). Changes in women's assertiveness in response to status and roles: A cross-temporal meta-analysis, 1931-1993. *Journal of Personality and Social Psychology*, 81, 133-145.
- Wang, L., Malhotra, D., & Murnighan, J. K. (2011). Economics education and greed. *Academy of Management Learning & Education*, 10, 643-660.
- The White House, Office of the Press Secretary. (2015, August 5). *Fact sheet: White House hosts convening of business school deans and the business community on expanding opportunities for women in business* [Press release]. Retrieved from <https://www.whitehouse.gov/the-press-office/2015/08/05/fact-sheet-white-house-hosts-convening-business-school-deans-and>
- Wood, W., Christensen, P. N., Hebl, M. R., & Rothgerber, H. (1997). Conformity to sex-typed norms, affect, and the self-concept. *Journal of Personality and Social Psychology*, 73, 523-535.
- Woodcock, A., Graziano, W. G., Branch, S. E., Habashi, M. M., Ngambeki, I., & Evangelou, D. (2013). Person and thing orientations: Psychological correlates and predictive utility. *Social Psychological and Personality Science*, 4, 116-123.
- Wright, R. E., & Bachrach, D. G. (2003). Testing for bias against female test takers of the Graduate Management Admissions Test and potential impact on admissions to graduate programs in business. *Journal of Education for Business*, 78, 324-328.
- 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.
- Yeaple, R. (2012, July 9). Is the MBA case method passé? *Forbes*. Retrieved from <http://www.forbes.com/sites/ronald-yeaple/2012/07/09/is-the-mba-case-method-passe/#1d486fea3145>