WOMEN IN SCIENCE

Expectations of brilliance underlie gender distributions across academic disciplines

Sarah-Jane Leslie,[1*†] Andrei Cimpian,[2*†] Meredith Meyer,[3] Edward Freeland[4]

The gender imbalance in STEM subjects dominates current debates about women's underrepresentation in academia. However, women are well represented at the Ph.D. level in some sciences and poorly represented in some humanities (e.g., in 2011, 54% of U.S. Ph.D.'s in molecular biology were women versus only 31% in philosophy). We hypothesize that, across the academic spectrum, women are underrepresented in fields whose practitioners believe that raw, innate talent is the main requirement for success, because women are stereotyped as not possessing such talent. This hypothesis extends to African Americans' underrepresentation as well, as this group is subject to similar stereotypes. Results from a nationwide survey of academics support our hypothesis (termed the field-specific ability beliefs hypothesis) over three competing hypotheses.

Laboratory, observational, and historical evidence reveals pervasive cultural associations linking men but not women with raw intellectual talent (*1–4*). Given these ambient stereotypes, women may be underrepresented in academic disciplines that are thought to require such inherent aptitude. We term this the field-specific ability beliefs hypothesis (fig. S1).

Current discourse about women in academia focuses mainly on women's underrepresentation in (natural) science, technology, engineering, and mathematics (STEM) (*5*). However, STEM disciplines vary in their female representation (fig. S2) (*5*, *6*). Recently, women have earned approximately half of all Ph.D.'s in molecular biology and neuroscience in the United States, but fewer than 20% of all Ph.D.'s in physics and computer science (*7*). The social sciences and humanities (SocSci/Hum) exhibit similar variability. Women are currently earning more than 70% of all Ph.D.'s in art history and psychology, but fewer than 35% of all Ph.D.'s in economics and philosophy (*7*). Thus, broadening the scope of inquiry beyond STEM fields might reveal new explanations and solutions for gender gaps (*8*). We offer evidence that the field-specific ability beliefs hypothesis can account for the distribution of gender gaps across the entire academic spectrum.

[1]Department of Philosophy, Princeton University, Princeton, NJ 08544, USA. [2]Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA. [3]Department of Psychology, Otterbein University, Westerville, OH 43081, USA. [4]Survey Research Center, Princeton University, Princeton, NJ 08544, USA.
***These authors contributed equally to the work.** †Corresponding author. E-mail: sjleslie@princeton.edu (S.-J.L.) or acimpian@illinois.edu (A.C.)

Individuals' beliefs about what is required for success in an activity vary in their emphasis on fixed, innate talent (*9*). Similarly, practitioners of different disciplines may vary in the extent to which they believe that success in their discipline requires such talent. Because women are often negatively stereotyped on this dimension (*1–4*), they may find the academic fields that emphasize such talent to be inhospitable. There are several mechanisms by which these field-specific ability beliefs might influence women's participation. The practitioners of disciplines that emphasize raw aptitude may doubt that women possess this sort of aptitude and may therefore exhibit biases against them (*10*). The emphasis on raw aptitude may activate the negative stereotypes in women's own minds, making them vulnerable to stereotype threat (*11*). If women internalize the stereotypes, they may also decide that these fields are not for them (*12*). As a result of these processes, women may be less represented in "brilliance-required" fields.

We used a large-scale, nationwide study of academics from 30 disciplines to evaluate the field-specific ability beliefs hypothesis, along with three competing hypotheses. The first competitor concerns possible gender differences in willingness or ability to work long hours (*13*): The more demanding a discipline in terms of work hours, the fewer the women. The second competing hypothesis concerns possible gender differences at the high end of the aptitude distribution [(*14*, *15*); but see (*16*, *17*) for criticism]. Such differences might cause greater gender gaps in fields that, by virtue of their selectivity, sample from the extreme right of the aptitude distribution: The more selective a discipline, the fewer the women. The third competing hypothesis concerns possible differences among fields in the extent to which they require systemizing (the ability to think systematically and abstractly) or empathizing (the ability to understand thoughts and emotions in an insightful way): The more a discipline prioritizes systemizing over empathizing, the fewer the women (*14, 18, 19*). Our findings suggest that the field-specific ability beliefs hypothesis, unlike these three competitors, is able to predict women's representation across all of academia, as well as the representation of other similarly stigmatized groups (e.g., African Americans).

We surveyed faculty, postdoctoral fellows, and graduate students (N = 1820) from 30 disciplines (12 STEM, 18 SocSci/Hum) (table S1) at geographically diverse high-profile public and private research universities across the United States. Participants were asked questions concerning their own discipline (table S2); responses in each discipline were averaged (tables S3 and S4), and analyses were conducted over disciplines (not individuals). As our dependent measure, we used the percentage of female Ph.D. recipients in each discipline (*7*).

To assess field-specific ability beliefs, we asked participants to rate their agreement with four statements concerning what is required for success in their field (e.g., "Being a top scholar of [discipline] requires a special aptitude that just can't be taught") (table S2). Respondents rated both the extent to which they personally agreed with these statements, and the extent to which they believed other people in their field would agree with the statements. Because answers to these eight questions displayed very similar patterns (α = 0.90), they were averaged to produce a field-specific ability belief score for each discipline (with higher scores indicating more emphasis on raw ability). As predicted, the more a field valued giftedness, the fewer the female Ph.D.'s. Field-specific ability belief scores were significantly correlated with female representation across all 30 fields [correlation coefficient $r(28) = -0.60$, $P < 0.001$], in STEM alone [$r(10) = -0.64$, $P = 0.025$], and in SocSci/Hum alone [$r(16) = -0.62$, $P = 0.006$] (Fig. 1). In a hierarchical regression with a STEM indicator variable entered in the first step and field-specific ability belief scores entered in the second (Table 1, models 1 and 2), adding the ability belief variable significantly increased the variance accounted for, $\Delta R^2 = 0.29$, $P < 0.001$.

To assess work demands, we asked participants to report the number of hours they worked per week, on-campus and off-campus (table S2). There was no correlation between the total number of hours worked (on- plus off-campus) and female representation, $r(28) = -0.03$, $P = 0.895$. Women tended to be underrepresented in fields whose practitioners worked more on-campus hours, but this correlation was not significant either, $r(28) = -0.32$, $P = 0.088$. No significant correlations with on-campus hours were found either within STEM, $r(10) = 0.46$, $P = 0.131$ (note the positive coefficient here), or within SocSci/Hum, $r(16) = -0.07$, $P = 0.772$. Adding on-campus hours to the hierarchical regression predicting

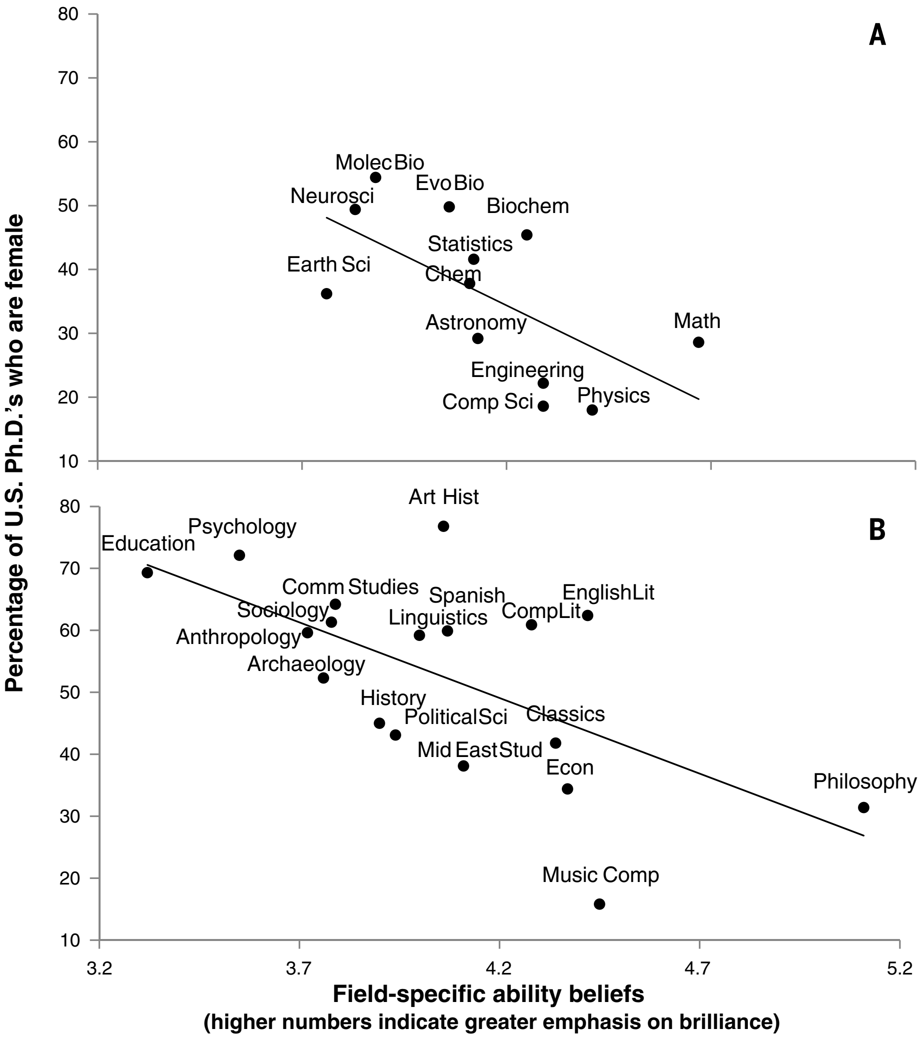

Fig. 1. Field-specific ability beliefs and the percentage of female 2011 U.S. Ph.D.'s in (A) STEM and (B) Social Science and Humanities.

women's representation did not significantly increase the variance accounted for, $\Delta R^2 < 0.01$, $P = 0.687$ (Table 1, model 3) [Similar results were obtained with total hours worked, as detailed in the supplementary materials (SM).] Thus, differences between fields in hours worked did not explain variance in the distribution of gender gaps beyond that explained by field-specific ability beliefs and the STEM indicator variable.

To assess selectivity, we asked faculty participants to estimate the percentage of graduate applicants admitted each year to their department. We then reverse-coded this measure so that higher values indicate more selectivity. Fields that were more selective tended to have higher, rather than lower, female representation, but this correlation did not reach significance, $r(28) = 0.34$, $P = 0.065$. Further, this selectivity measure did not predict female representation in STEM alone or in SocSci/Hum alone (both *P*s > 0.478), and adding it to the hierarchical regression did not result in a statistically significant increase in the variance accounted for, $\Delta R^2 = 0.04$, $P = 0.134$ (Table 1, model 4). (An analysis considering only selectivity measures from top-10% departments produced the same pattern of results; see the SM.) To account for potential differences in the strength of the applicant pools across disciplines, we compared the 2011–2012 Graduate Record Examination (GRE) General Test scores of Ph.D. applicants. These data were available for only 19 of the disciplines in our study (7 STEM and 12 SocSci/Hum) (*20*). A composite measure of GRE scores was not significantly correlated with female representation, $r(17) = -0.24$, $P = 0.333$, and so provided no evidence that fields with more women have weaker applicant pools. Further, the relation between field-specific ability beliefs and female representation remained significant when adjusting for GRE scores, $r(16) = -0.57$, $P = 0.013$.

Table 1. Hierarchical regression models predicting female representation. $N = 30$ disciplines. Significant statistics are bold. R^2 comparisons are always with the preceding model (to the left).

Predictor	Model 1			Model 2			Model 3			Model 4			Model 5		
	β	*t*	*P*	β	*t*	*P*	β	*t*	*P*	β	*t*	*P*	β	*t*	*P*
STEM indicator	**–0.50****	**–3.03**	**0.005**	**–0.42****	**–3.20**	**0.003**	–0.35	–1.49	0.148	–0.30	–1.34	0.193	–0.28	–1.07	0.297
Field-specific ability beliefs				**–0.55*****	**–4.13**	**<0.001**	**–0.56*****	**–3.98**	**<0.001**	**–0.58*****	**–4.17**	**<0.001**	**–0.56****	**–3.46**	**0.002**
On-campus hours worked							–0.09	–0.41	0.687	–0.01	–0.03	0.975	0.02	0.07	0.945
Selectivity										0.24	1.55	0.134	0.24	1.54	0.137
Systemizing versus empathizing													–0.06	–0.23	0.817
R^2		**0.25**			**0.54**			0.54			0.58			0.58	
F for change in R^2		**9.19****			**17.08*****			0.17			2.40			0.06	
P for change in R^2		**0.005**			**<0.001**			0.687			0.134			0.817	

***P*< 0 0.01. ****P* < 0.001.

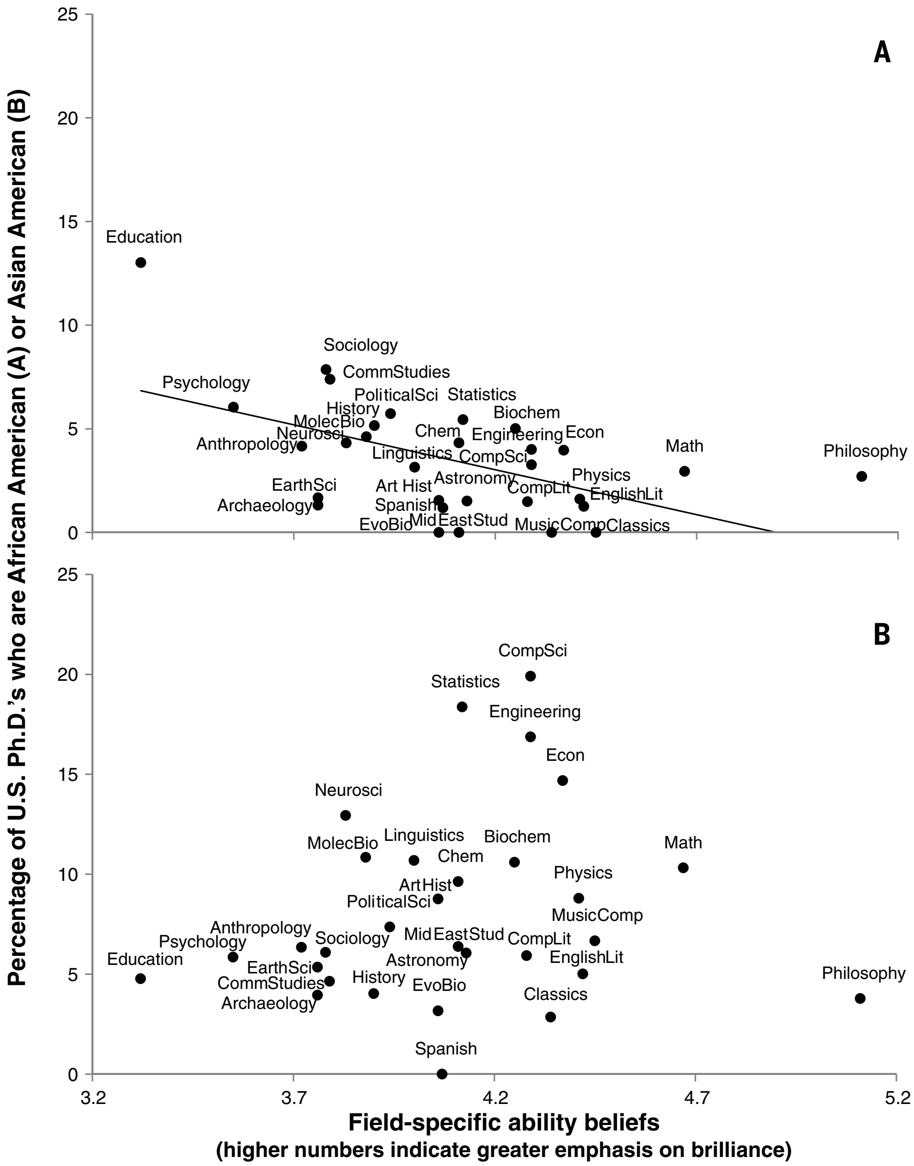

Fig. 2. Field-specific ability beliefs and the percentage of 2011 U.S. Ph.D.'s who are (A) African American and (B) Asian American.

To assess systemizing versus empathizing, we asked participants to evaluate the extent to which scholarly work in their discipline requires these two abilities (two questions each, table S2) (αs = 0.63 and 0.90, respectively). A composite systemizing-minus-empathizing score was significantly correlated with female representation across all disciplines, $r(28) = -0.53$, $P = 0.003$. However, this score did not significantly predict female representation in STEM alone, $r(10) = -0.27$, $P = 0.402$, or in SocSci/Hum alone, $r(16) = -0.25$, $P = 0.310$. Adding systemizing versus empathizing composite scores to the hierarchical regression did not increase the variance accounted for, $\Delta R^2 < 0.01$, $P = 0.817$ (Table 1, model 5). Indeed, field-specific ability beliefs were the sole significant predictor of female representation in this final model, $\beta = -0.56$, $P = 0.002$.

To further compare our hypothesis to its competitors, we performed another hierarchical regression in which the STEM indicator, on-campus hours, selectivity, and systemizing versus empathizing were all added together as a first step. The model was significant, $R^2 = 0.38$, $P = 0.016$, although no individual predictor in it was. When field-specific ability belief scores were added, the variance accounted for increased to 58%, $\Delta R^2 = 0.21$, $P = 0.002$. This finding reflects once again the predictive power of the field-specific ability beliefs hypothesis. (However, we do not claim that field-specific ability beliefs are the sole determiner of gender gaps or that these three are the only alternative hypotheses; other factors undoubtedly play a role. Further, this study was not designed to eliminate the three competing hypotheses but rather to use them as a benchmark for our hypothesis.)

To check for robustness, we duplicated our analyses using weights created by comparing the demographic characteristics of our respondents against the demographics of the population of academics we initially contacted. Adding these post-stratification weights to our analyses helped us to check for the influence of bias resulting from differential nonresponse (*21*, *22*) (for details, see the SM). In a weighted version of the final model in the hierarchical regression, field-specific ability beliefs were again the only significant predictor of female representation, $\beta = -0.40$, $P = 0.029$ (see table S7 for weighted versions of all models).

There are many potential mechanisms by which field-specific ability beliefs may influence women's representation. To assess some possibilities, we asked participants to evaluate the statement, "Even though it's not politically correct to say it, men are often more suited than women to do high-level work in [discipline]." Participants rated their own agreement and the extent to which they thought that other people in their field would agree. These two scores were averaged into a single measure (α = 0.80). Disciplines that emphasized raw talent were more likely to endorse the idea that women are less suited for high-level scholarly work, $r(28) = 0.38$, $P = 0.036$. In turn, higher endorsement of this idea was associated with lower female representation, $r(28) = -0.67$, $P < 0.001$. We also asked participants to rate whether they thought that their discipline was welcoming to women (table S2). Disciplines that valued giftedness over dedication rated themselves as less welcoming to women, $r(28) = -0.58$, $P = 0.001$, and fields that viewed themselves as less welcoming had fewer female Ph.D.'s, $r(28) = 0.74$, $P < 0.001$. Together, ratings of whether women were suitable for and welcome in a discipline mediated 70.2% of the relation between field-specific ability beliefs and the percentage of female Ph.D.'s (bootstrapped $P < 0.001$) (*23*). Thus, field-specific ability beliefs may lower women's representation at least in part by fostering the belief that women are less well-suited than men to be leading scholars and by making the atmosphere in these fields less welcoming to women.

Like women, African Americans are stereotyped as lacking innate intellectual talent (*24*). Thus, field-specific ability belief scores should predict the representation of African Americans across academia. Indeed, African Americans were less well represented in disciplines that believed giftedness was essential for success, $r(28) = -0.54$, $P = 0.002$ (Fig. 2). However, field-specific ability belief scores should not predict the representation of Asian Americans, who are not stereotyped in the same way (*25*). Indeed, Asian American representation was not correlated with field-specific ability beliefs, $r(28) = 0.16$, $P = 0.386$. A hierarchical regression with the STEM indicator, total hours worked (the same results are found with on-campus hours only), selectivity, and systemizing versus empathizing, all entered together in

Table 2. Hierarchical regression models predicting African American representation. N = 30 disciplines. Significant statistics are bold. R^2 comparisons are always with the preceding model (to the left).

Predictor	Model 1			Model 2		
	β	*t*	*P*	β	*t*	*P*
STEM indicator	0.05	0.15	0.878	−0.06	−0.21	0.833
Total hours worked	−0.13	−0.58	0.571	−0.26	−1.30	0.207
Selectivity	−0.21	−0.88	0.387	−0.19	−0.93	0.359
Systemizing versus empathizing	−0.23	−0.71	0.487	0.10	0.35	0.733
Field-specific ability beliefs				**−0.61****	**−3.28**	**0.003**
R^2		0.06			**0.35**	
F for change in R^2		0.37			**10.76****	
P for change in R^2		0.827			**0.003**	

**P < 0.01.

the first step predicted very little variance in African Americans' representation, $R^2 = 0.06$, P = 0.827 (Table 2, model 1). Adding field-specific ability beliefs in the second step significantly increased the variance accounted for, $\Delta R^2 = 0.29$, P = 0.003 (Table 2, model 2). As with women, field-specific ability beliefs were the only significant predictor of African American representation (with total hours worked: β = −0.61, P = 0.003; with on-campus hours: β = −0.64, P = 0.005).

Finally, we considered alternative explanations for our results. If women believe more strongly than men in the value of hard work, disciplines with fewer women may have higher field-specific ability belief scores precisely for that reason—because they have fewer women. Relative to men, women in our survey did report lower field-specific ability belief scores and thus more belief in the importance of dedication ($M_{\text{women}} = 3.86$ versus $M_{\text{men}} = 4.24$), $t(1812) = 7.31$, $P < 0.001$. However, contrary to this alternative hypothesis, the field-specific ability belief scores derived separately from each discipline's female and male respondents were independently predictive of the percentage of female Ph.D.'s across the 30 fields: $r(28) = -0.40$, P = 0.028, for women's scores, and $r(28) = -0.50$, P = 0.005, for men's. We also constructed a gender-balanced field-specific ability belief score for each discipline by computing the average scores for men and women in that discipline and then averaging the two gender-specific scores. This measure weights women's and men's scores equally, regardless of their actual representation in a field, and was again the only variable that significantly predicted women's representation in the regression model that included all competitors plus the STEM indicator, β = −0.49, P = 0.009. Thus, the relation between field-specific ability beliefs and women's representation is not a simple matter of men and women valuing effort differently.

Is it possible that people simply infer what is required for success in a field on the basis of their estimates of that field's diversity, so that they assume a field requires less brilliance if women are well represented in it? Contrary to this alternative hypothesis, a regression analysis performed on the individual-level data revealed that academics' perceptions of the diversity of their field (see table S2) were in fact not a significant predictor of their field-specific ability beliefs, β = −0.05, P = 0.277. This analysis included indicator variables for each discipline so as to minimize the influence of discipline-specific unobserved variables (*26*). Thus, the relation between field-specific ability beliefs and women's representation is not a simple matter of using women's representation in a field to infer what is required for success.

Is natural brilliance truly more important to success in some fields than others? The data presented here are silent on this question. However, even if a field's beliefs about the importance of brilliance were to some extent true, they may still discourage participation among members of groups that are currently stereotyped as not having this sort of brilliance. As a result, fields that wished to increase their diversity may nonetheless need to adjust their achievement messages.

Are women and African Americans less likely to have the natural brilliance that some fields believe is required for top-level success? Although some have argued that this is so, our assessment of the literature is that the case has not been made that either group is less likely to possess innate intellectual talent (as opposed to facing stereotype threat, discrimination, and other such obstacles) (*10*, *16*, *17*, *24*, *27*).

The extent to which practitioners of a discipline believe that success depends on sheer brilliance is a strong predictor of women's and African Americans' representation in that discipline. Our data suggest that academics who wish to diversify their fields might want to downplay talk of innate intellectual giftedness and instead highlight the importance of sustained effort for top-level success in their field. We expect that such easily implementable changes would enhance the diversity of many academic fields.

REFERENCES AND NOTES

1. M. Bennett, *J. Soc. Psychol.* **136**, 411–412 (1996).
2. B. Kirkcaldy, P. Noack, A. Furnham, G. Siefen, *Eur. Psychol.* **12**, 173–180 (2007).
3. A. Lecklider, *Inventing the Egghead* (Univ. of Pennsylvania Press, Philadelphia, 2013).
4. J. Tiedemann, *Educ. Stud. Math.* **41**, 191–207 (2000).
5. S. J. Ceci, W. M. Williams, *Why Aren't More Women in Science?* (APA Books, Washington, DC, 2007).
6. S. J. Ceci, W. M. Williams, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3157–3162 (2011).
7. National Science Foundation, Survey of Earned Doctorates, (2011); www.nsf.gov/statistics/srvydoctorates/.
8. S. J. Ceci, D. K. Ginther, S. Kahn, W. M. Williams, *Psychol. Sci. Public Interest* **15**, 75(2014).
9. C. S. Dweck, *Mindset: The New Psychology of Success* (Random House, New York, 2006).
10. V. Valian, *Why So Slow? The Advancement of Women* (MIT Press, Cambridge, MA, 1998).
11. I. Dar-Nimrod, S. J. Heine, *Science* **314**, 435 (2006).
12. A. Wigfield, J. S. Eccles, *Contemp. Educ. Psychol.* **25**, 68–81 (2000).
13. K. Ferriman, D. Lubinski, C. P. Benbow, *J. Pers. Soc. Psychol.* **97**, 517–532 (2009).
14. D. C. Geary, *Male, Female: The Evolution of Human Sex Differences* (APA Books, Washington, DC, ed. 2, 2010).
15. L. V. Hedges, A. Nowell, *Science* **269**, 41–45 (1995).
16. J. S. Hyde, *Am. Psychol.* **60**, 581–592 (2005).
17. A. M. Penner, *Am. J. Sociol.* **114** (suppl.), S138–S170 (2008).
18. J. Billington, S. Baron-Cohen, S. Wheelwright, *Learn. Individ. Differ.* **17**, 260–268 (2007).
19. R. Lippa, *J. Pers. Soc. Psychol.* **74**, 996–1009 (1998).
20. Educational Testing Service, *GRE: Guide to the Use of Scores* (ETS, Princeton, NJ, 2012).
21. J. Bethlehem, *Int. Stat. Rev.* **78**, 161–188 (2010).
22. R. J. A. Little, *J. Am. Stat. Assoc.* **88**, 1001–1012 (1993).
23. A. F. Hayes, *Introduction to Mediation, Moderation, and Conditional Process Analysis* (Guilford Press, New York, 2013).
24. C. M. Steele, J. Aronson, *J. Pers. Soc. Psychol.* **69**, 797–811 (1995).
25. M. Shih, T. L. Pittinsky, N. Ambady, *Psychol. Sci.* **10**, 80–83 (1999).
26. J. D. Angrist, J. S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton Univ. Press, Princeton, NJ, 2008).
27. S. J. Gould, *The Mismeasure of Man* (Norton, New York, 1996).

ACKNOWLEDGMENTS

We thank E. Anastasia, R. Baillargeon, R. Bigler, C. Chaudhry, R. Fusello, E. Harman, M. Johnston, Y. Lee, S. Lubienski, A. Penner, H. Perhai, E. Pomerantz, J. Robinson-Cimpian, and the members of the Cognitive Development Lab, University of Illinois at Urbana-Champaign. We also thank five anonymous reviewers for their comments. This work was supported by NSF grant BCS-1226942 to S.-J.L. and by Spencer Foundation grant 201100111 to A.C. The data are available in the SM, with values for the primary variables given in table S4. The SM contains additional data.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6219/262/suppl/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S8
References

17 September 2014; accepted 25 November 2014
10.1126/science.1261375