Title
Can Short Psychological Interventions Affect Educational Performance? Revisiting the Effect of Self-Affirmation Interventions

Journal
Psychological Science, 31(7)

ISSN
0956-7976

Authors
Serra-Garcia, Marta
Hansen, Karsten T
Gneezy, Uri

Publication Date
2020-07-01

DOI
10.1177/0956797620923587

Can Short Psychological Interventions Affect Educational Performance? Revisiting the Effect of Self-Affirmation Interventions

Authors: M. Serra-Garcia[1]*, K. Hansen[1], U. Gneezy[1,2],

Affiliations:

[1] UC San Diego, Rady School of Management
[2] CREED, University of Amsterdam
*Correspondence to: mserragarcia@ucsd.edu.

Abstract: Large amounts of resources are spent annually to improve educational achievement and to close the gender gap in sciences with typically very modest effects. Miyake et al., (*Science*, 2010) introduced a 15-minute "self-affirmation" intervention and reported a dramatic reduction in the gender gap. This paper re-analyzes the original data and finds several critical problems. First, the self-affirmation hypothesis states that women's performance will improve. However, the data showed no improvement for women. There was an interaction effect between self-affirmation and gender, caused by a negative effect on men's performance. Second, the findings were based on covariate-adjusted interaction effects, which imply that self-affirmation only reduced the gender gap for the small sample of men and women who did not differ in the covariates. Third, specification-curve analyses with over 1,500 possible specifications showed that fewer than a quarter yielded significant interaction effects and fewer than three percent showed significant improvements among women.

Keywords: self-affirmation, gender gap, education.

Education in the United States is far from reaching desired outcomes (Swanson, 2009). The problem is not in the number of policy ideas that are tried; it is in the success of these policies. For example, in a large-scale field experiment to study whether monetary incentives can improve educational outcomes, some students were paid to read books, others for performance on interim assessments (Fryer, 2011). Despite spending $9.4 million, the impacts were statistically indistinguishable from zero (see also, Springer et al., 2009, Gneezy et al., 2011, Fryer et al., 2012).

A particular focus of many programs is to reduce the gender gap in science, technology, engineering, and mathematics (STEM). Given the limited effectiveness and high cost of most interventions, the success of a very simple psychological manipulation tested by Miyake et al. (2010) is impressive. In the intervention, called *values affirmation*, 399 students in a college-level introductory physics class either wrote about their most important values or their least important values (in the control condition), twice for 15 minutes within the first four weeks of the 15-week course.

Miyake et al. (2010) reported that the intervention reduced the gender gap in physics exams at the 1% significance level. While studies using monetary incentives to improve performance find average effects ranging from –0.031 to 0.079 standard deviations (SD) (Fryer, 2011), the effects in Miyake et al. (2010) ranged from 0.12 to 0.19 SD. Similar strong results of the self-affirmation intervention have been found in schools (Cohen et al., 2006), even in two-year and 7-to-9-year follow up studies (Cohen et al., 2009; Goyer et al., 2017).

Miyake et al. (2010) is a widely cited study that has had a substantial impact on the field of STEM education (Nisbett, 2010; Hanselman et al., 2017) and the broader use of self-affirmation interventions. However, other studies (Borman, 2012; Lauer et al., 2013; Dee, 2015; De Jong et al., 2016; Bratter, Rowley and Chukhray, 2016; Hanselman et al., 2017; Hoffman and Kurtz-Cortes 2019) failed to replicate the findings.

We re-analyzed the Miyake et al. (2010) data, and examined it at a conceptual and a statistical level. At the conceptual level, value affirmation theory starts by observing the common stereotype that men are better at math and science than women. This may generate increased academic pressure for women who subscribe to this stereotype and possibly negatively affect their academic performance. To offset this, the theory proposes that a values affirmation task, in which people reflect on self-defining values, can buffer people against such psychological threat. The resulting hypothesis is that (some) women will improve as a result of the intervention. This theory has no prediction for male students. However, the original data in Miyake et al. (2010) showed that the observed reduction in the gender gap was a result of the interaction effect of the intervention on the performance of men and women. We examined the results for male and female students separately.

From a statistical perspective, the analysis in Miyake et al. (2010) was based on covariate-adjusted means. The interaction effect of values affirmation must be thus interpreted as conditional on a *given level of prior performance and stereotype endorsement* (Cochran, 1957; Miller and Chapman, 2001). However, since prior performance and stereotype endorsement on average differed substantially for men and women, the reported effect was restricted to a small portion (28%) of the sample: those females with the same performance and stereotype endorsement as males.

Miyake et al. (2010) reported the results of one empirical model specification. Using specification-curve analyses (Simonsohn et al., 2015; see also, Steegen et al., 2016), we examined the robustness of the effect of values affirmation on the gender gap and on females' performance.

Finally, Miyake et al. (2010) argued that results were consistent with theoretical accounts of stereotype threat because the intervention improved the performance of women displaying high stereotype endorsement. To show this, they compared women whose stereotype endorsement is either 0.75 SD higher than the mean (high endorsement) or 0.75 SD lower than the mean (low endorsement). We re-examined these heterogeneous effects considering the responses of all women along the 5-point scale measuring stereotype endorsement.

We also ran a small replication study of Miyake et al. (2010) in a physics class at UC San Diego, because replication is an important part of social science (Simmons et al., 2011; Open Science Collaboration, 2012, 2015; Simonsohn, 2016; Camerer et al., 2016), and report the results of our suggestive study in the SOM.

Method

Revisiting Miyake et al. (2010)

The self-affirmation intervention worked as follows. Students in the treated group wrote about personally important values (such as friends and family). Students in the control group selected their least important values from the same list and wrote why these values might be important to other people. Thus, all students wrote about values and their importance, but the exercise was self-relevant only for the affirmation group (the students in the treated group). In Miyake et al. (2010) this 15-minute writing exercise was conducted during class once in week one and once in an online homework assignment during week four of the semester of an introductory physics class in college. This brief exercise was not related to the subject matter of the course.

The theory is that "Values affirmation, in which people reflect on self-defining values, can buffer people against such psychological threat. When they affirm their core values in a threatening environment, people reestablish a perception of personal integrity and worth, which in turn can provide them with the internal resources needed for coping effectively…", according to Miyake et al. (2010). The theory does not predict a negative effect on men but only a positive one for women.

We solicited the raw data from the original study by Miyake et al. (2010) to better understand their results. These data included three continuous measures of performance: mean exam scores, scores on a standardized physics test (the Force and Motion Concept Evaluation or FMCE), and the final course score, based substantially (75%) on the exam scores.

Results

We found that, using the original data by Miyake et al. (2010), values affirmation had no significant effect on average female performance (covariate-unadjusted). If at all, it had a negative effect on male performance in some dimensions. Covariate-unadjusted means are reported in Table S1 in the supplementary materials of Miyake et al. (2010).

Values affirmation had no significant effect on the mean exam score of females (Cohen's d=0.19) ($t(114)$=1.00, p=0.318). However, the performance of male students was significantly lower in the values affirmation condition relative to the control condition

(Cohen's *d*=-0.25) ($t(281)=-2.06$, $p=0.040$). The interaction effect was significant ($\beta_i=0.43$, $t(395)=1.97$, $p=0.050$), but it was driven by a drop in male performance.

Values affirmation did not have a significant effect on male or female scores on a standardized physics test (FMCE score) at the end of the semester. The scores of female students did not differ significantly across the values affirmation and the control condition (Cohen's *d*=0.31) ($t(94)=1.28$, $p=0.21$). Male students did not obtain significantly different scores across conditions either (Cohen's *d*=-0.09) ($t(210)=-0.54$, $p=0.587$). The interaction effect between values affirmation and gender was not significant ($\beta_i=0.34$, $t(304)=1.40$, $p=0.164$).

Fig. 1 presents the results obtained by Miyake et al. (2010) for final course scores. The final course scores obtained by females and males were not affected by the values affirmation intervention, and the interaction effect was not significant either. All details of the regression analysis are reported in the SOM.

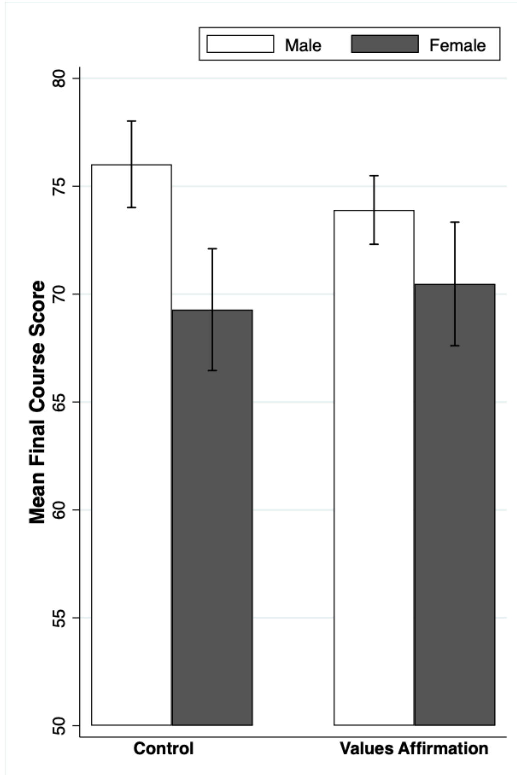


Fig. 1: Mean final course scores by condition and gender in Miyake et al (2010). Confidence interval bars are indicated.

The interpretation of covariate-adjusted effects

These results contrast with the covariate-adjusted results reported by Miyake et al. (2010). Miyake et al. (2010) reported a significant positive effect of the values affirmation intervention on the performance of females using covariate-adjusted means. The adjustment was based on two covariates: (1) prior performance, measured by the SAT score for the mean exam score and the course score, and the beginning-of-semester FMCE score for the FMCE score at the end of the semester; (2) and stereotype endorsement, measured by students' agreement with the statement "According to my own personal beliefs, I expect men to generally do better in physics than women" on a 5-point scale ranging from "strongly disagree" to "strongly agree." The authors argued that it is critical to assess the effects of the self-affirmation intervention "controlling for prior relevant performance" (Supplementary Materials, p. 13, Miyake et al. (2010)) and that background variables, including stereotype endorsement, should be included and interacted with the experimental difference variables. They pooled male and female students together and examined the effect of the self-affirmation intervention, including variables for gender, stereotype endorsement and prior performance, as well as interaction terms between these.

The interpretation of covariate-adjusted effects is that there was a reduction in the gender gap *for a population of men and women who have the same SAT score* (and level of stereotype endorsement). We demonstrate the interpretation of covariate-adjusted effects formally in the SOM. Is the covariate-adjusted effect the effect we are interested in? It could be if the distributions of SAT scores and stereotype endorsement for men and women had similar means. However, there was a significant difference in prior performance of men and women (for SAT scores, $t(397)=2.62$, $p=0.01$; for beginning-of-semester FMCE scores, $t(306)=4.80$, $p<0.01$) and in stereotype endorsement ($\chi2(4)=41.64$ $p<0.01$). Hence, the effect that was estimated with covariate adjustment was only relevant for the subset of women who had the same prior performance and stereotype endorsement as men. In Miyake et al. (2010), this was 56% of the sample, considering only SAT or ACT scores. Including stereotype endorsement, only 28% of the sample featured male and female students with the same SAT scores and stereotype endorsement.

The relevant question is whether values affirmation had a significantly positive effect on average *female* performance. We examined this effect as part of the specification-curve analysis described next.

Evaluating Analytical Approaches: Specification-Curve Analysis

As we have discussed above, adding covariates such as SAT scores, and their interactions, changes the interpretation of the coefficients, and also invalidates standard linear model analysis. The researcher may be interested in understanding how robust the effect of self-affirmation interventions is to the inclusion of different covariates in the regression model, as well as other decisions regarding the data analysis (e..g, Steegen et al., 2016; Simonsohn et al., 2015). To investigate this, we conducted a specification curve analysis (Simonsohn et al., 2015).

The regression model used in Miyake et al. (2010) included 11 independent variables. In addition to gender (F_i), affirmation condition (Z_i) and prior performance (S_i), it included stereotype threat, which we denote as $T_{i.}$ The specification included 2-way and 3-way interaction effects of these variables, and was as follows:

$$Y_i \;=\; \beta + \beta_Z \times Z_i + \beta_F \times F_i + \beta_{FZ} \times F_i \times Z_i + \beta_S \times S_i + \beta_{SZ} \times S_i \times Z_i + \beta_{SF} \times S_i \times F_i$$
$$\beta_T \times T_i + \beta_{TZ} \times T_i \times Z_i + \beta_{TF} \times T_i \times F_i + \beta_{TS} \times T_i \times S_i + \beta_{TZF} \times T_i \times Z_i \times F_i + \varepsilon_i,$$

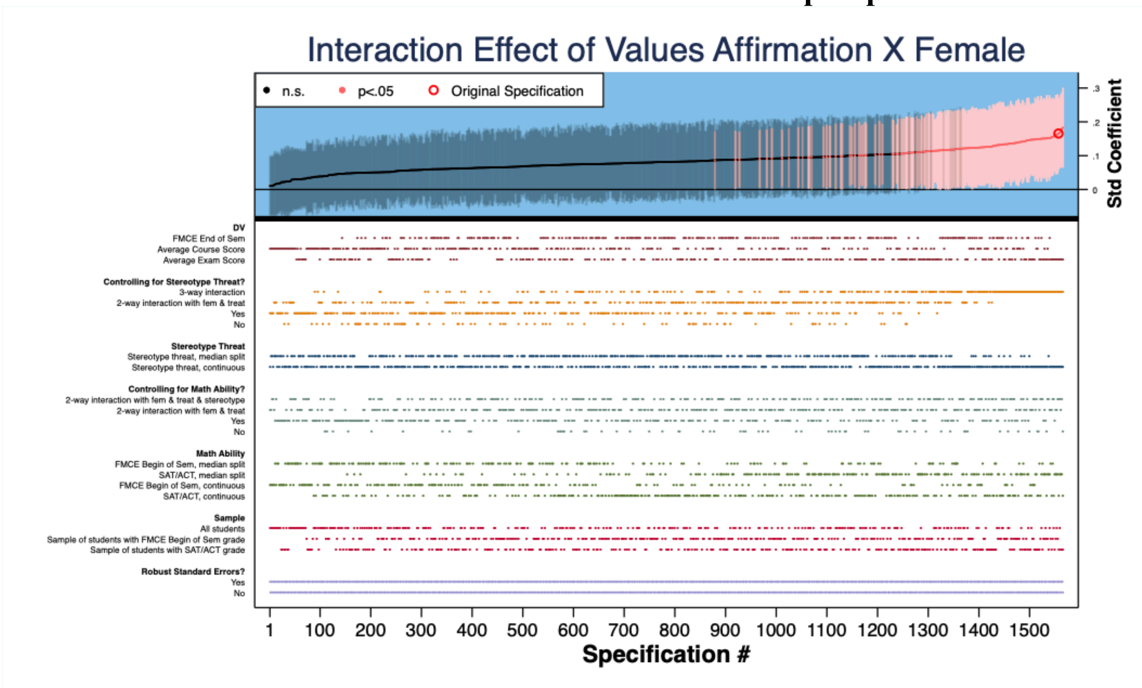
Miyake et al. (2010) focused on the average exam score as their main outcome of interest, but also included final course score and score in the FMCE at the end of the semester in some of the analyses. We investigated how the gender gap in academic performance (values affirmation X gender interaction) changed under the following alternatives: (1) using the three different dependent variables mentioned above, all of which measure academic performance; (2) including stereotype threat as a covariate, interacted or not with other covariates; (3) using different definitions of stereotype threat (as a continuous variable, or a dummy for those above median); (4) including prior performance, interacted or not with other covariates; (5) using different variables and definitions of prior performance (SAT/ACT score or FMCE score at the beginning at the semester, continuous or as a dummy for those performing above median); (6) allowing for different exclusions of missing observations; (7) and estimating robust standard errors. These decisions yielded a total of 1566 unique estimated interaction effects of values affirmation and gender. Further details of the analysis and results are provided in the SOM. Since the relevant question is whether the values affirmation has a significantly positive effect on average *female* performance, we also conducted the specification curve analysis for the effect of values affirmation on female students.

Panel A of Figure 2 plots the coefficient and confidence interval of the interaction of values affirmation and gender at the sample average. The dependent variable was standardized, such that the coefficient can be interpreted in standard deviations. Panel B of Figure 2 plots the coefficient and confidence interval of the effect of values affirmation on female students. Each figure also shows the distribution of effects for each of the researcher's decisions, described on the left-hand side.

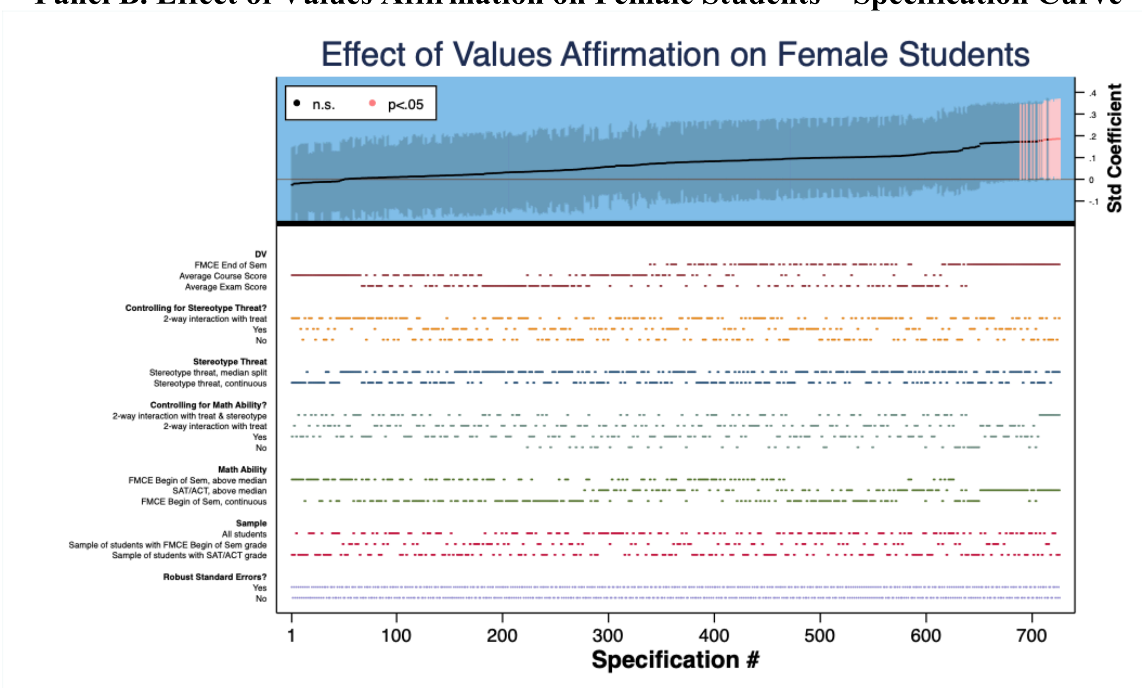
We replicated the *t*-statistic for the interaction effect in the original work of Miyake et al. (2010), 3.08, and indicate where it lies on the specification curve. It was the 15th highest out of 1566 specifications. The interaction effect was not statistically significant ($p>0.05$) in 1205 specifications. That is, in 77% of specifications the effect of values affirmation was not different between female and male students. As shown in Figure 2, a key decision in obtaining a significant interaction effect was to include the 3-way interaction between values affirmation, gender and stereotype threat endorsement. If it was not included, 91% of specifications (1000 out of 1098) were not significant. By contrast, of the 361 significant interaction effects, 73% include the 3-way interaction. This indicates that, not only was covariate-adjustment needed, but that a 3-way interaction term was needed to find a significant reduction of the gender gap in academic performance.

Within the group of female students, the treatment was assigned randomly and females in the treatment and control conditions do not vary in their prior performance (for SAT scores by condition, $t(114)=-0.79$, $p=0.4329$; for beginning-of-semester FMCE score by condition, $t(94)=-0.65$, $p=0.5187$) or stereotype endorsement ($\chi^2(4)=2.46$, $p=0.653$). Hence, covariate adjustment could be used to reduce variance (Miller and Chapman, 2001) without reducing the effect of values affirmation to a specific group of students. In the data shared with us by Miyake et al. (2010), SAT/ACT grades were standardized for men and women and, with the data available, we could not recode them and use them within the sample of women. Stereotype endorsement and FMCE scores at the beginning of the semester were sample centered. They could be recoded and we used these as covariates within the sample of women. The specification-curve is shown in Panel B of Figure 2. The effect of values affirmation on female students was not significant in 97.0% of specifications.

Panel A. Effect of Values Affirmation on the Gender Gap - Specification Curve



Panel B. Effect of Values Affirmation on Female Students – Specification Curve



NOTES: (a) Std Coefficient is the average interaction effect of values affirmation and gender (Panel A) or values affirmation on female students' performance (Panel B), where performance is always measured on the standardized dependent variable (DV). (b) To conduct the analysis in Panel B, SAT/ACT grade cannot be used because it is standardized for male and female students in the original data and its raw value cannot be recovered.

Figure 2. Specification Curve Analysis

Heterogeneity of Values Affirmation Effects by Stereotype Endorsement

Miyake et al. (2010) argued that "values affirmation was particularly beneficial for women who tended to endorse the gender stereotype" (p. 1236). They split the sample between high and low stereotype endorsement based on whether the student was +/- 0.75 SD from the mean. We examined this finding in further detail to understand the role of this particular sample split. Stereotype endorsement was based on agreement with the statement "According to my own personal beliefs, I expect men to generally do better in physics than women," with responses ranging from strongly disagree to strongly agree on a 5-point scale. Out of 96 female students without missing information, 7 (4 treated, 3 control) agreed with the stereotype, 10 (4 treated, 6 control) neither agreed nor disagreed, 24 disagreed (14 treated, 10 control) and 55 (33 treated, 22 control) strongly disagreed. Values affirmation had a positive and significant effect on the exam scores of women who agreed with the stereotype (4 women) and for women who neither agreed nor disagreed (a total of 8 women). Surprisingly, for women who strongly disagreed with the statement we observed negative effects of values affirmation on their exam score (33 women).

Using a regression analysis of performance on the female sample, controlling for stereotype endorsement and prior performance, the interaction between Values Affirmation and Stereotype Endorsement was significantly positive. Yet, this only indicates that the effect of values affirmation was *more* positive on the students with higher stereotype endorsement. It should not be interpreted as being a positive effect for everyone, and more strongly positive for women. As described above, the effect was negative for a large group of students. Detailed results are presented in the SOM.

Discussion

Understanding how to improve academic performance, in particular for struggling students, is an important challenge for social scientists and policy makers. Miyake et al. (2010) offered a remarkably strong and cost-effective way of doing that. Our investigation was motivated by a desire to better understand the psychological instrument in Miyake et al. (2010), and hopefully use it in large-scale interventions. A closer look at the data of Miyake et al. (2010) revealed two critical problems. The self-affirmation hypothesis only predicted an effect on women. However, their conclusions were based on the significant interaction effect between the self-affirmation intervention and gender. This effect was driven by a reduction in the performance of men, and not by an improvement in the performance of women. Further, the conclusions from the original study were highly sensitive to the empirical specification. Miyake et al. (2010) wrote that "values affirmation is a promising intervention that can help reduce the gender achievement gap in physics" (p. 1237). The statistical analysis presented here suggests that this conclusion was not supported by the data. Taking the evidence together, the results were supportive of a null effect. It is important to carefully understand the effects of such interventions because promoting ineffective interventions is costly in terms of resource allocation and negatively affects the success of policy makers' attempts to reduce the gender achievement gap.

Author Contributions

U. Gneezy and M. Serra-Garcia developed the study concept. K. Hansen analyzed the regression models in Miyake et al. (2010) and provided the interpretation of covariate-

adjusted effects. M. Serra-Garcia conducted the reanalysis of the data in Miyake et al. (2010). All authors drafted and revised the manuscript.

Acknowledgments

We thank Tiffany Ito for sharing the materials and data, and Anna Dreber, Leif Nelson, and Johannes Haushofer for valuable comments.

Declaration of Conflicting Interests

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

Supplemental Material

This project was conducted under IRB #130701XX. Additional supporting information is provided in the SOM.

Open Practices

STATA codes used for different analyses are available from the corresponding author upon request. Data was obtained from the authors of the original study (Miyake et al., 2010).

References:

1. Swanson, C.B., Editorial Projects in Education. (2009)

2. Fryer, R.G. Jr. Financial incentives and student achievement: Evidence from randomized trials. *Q. J. Econ*. 126 (4), 1755 (2011).
3. Springer, M.G., Ballou, D., Hamilton, L.S., Le, V., Lockwood, J.R., McCaffrey, D.F., Pepper, M., Stecher, B. Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching (POINT). Conference paper, National Center on Performance Incentives. (2010).
4. Gneezy, U., Meier, S., Rey-Biel, P. When and Why incentives (Don't) Work to Modify Behavior. *J of Economic Perspectives*, 25(4), 191-209 (2011).
5. Fryer, R.G. Jr, Levitt, S.D., List, J., Sadoff, S. Enhancing the efficacy of teacher incentives through loss aversion: A field experiment. NBER Working Paper 18237 (2012).
6. Miyake, A., Kost-Smith, L.E., Finkelstein, N.D., Pollock, S.J., Cohen, G.L., Ito, T.A. Reducing the Gender Achievement Gap in College Science: A Classroom Study of Values Affirmation. *Science* 330, 1234 (2010).
7. G. L. Cohen, J. Garcia, N. Apfel, A. Master, Reducing the racial achievement gap: A social-psychological intervention. *Science* 313, 1307 (2006).
8. G. L. Cohen, J. Garcia, V. Purdie-Vaughns, N. Apfel, P. Brzustoski, Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science* 324, 400 (2009).
9. J. P. Goyer, J. Garcia, V. Purdie-Vaughns, K.R. Binning, J.E. Cook, S.L. Reeves, N. Apfel, S. Taborsky-Barba, D.K. Sherman, G.L. Cohen. Self-affirmation facilitates

minority middle schoolers' progress along college trajectories. *Proceedings of the National Academy of Sciences* 114 (29), 7594-7599 (2017).

10. Nisbett, R.E. Think Big, Bigger … and Smaller. *Educational Leadership* 68 (3), Closing Opportunity Gaps, 10-15 (2010).
11. Hanselman, P., Rozek, C.S., Grigg, J., and Borman, G.D. New Evidence on Self-Affirmation Effects and Theorized Sources of Heterogeneity from Large-Scale Replications. *Journal of Educational Psychology* 109 (3), 405-424 (2017).
12. Lauer, S., Momsen, J., Offerdahl, E., Kryjevskaia, M., Christensen, W., & Montplaisir, L. Stereotyped: Investigating gender in introductory science courses. *CBE-Life Sciences Education,* 12, 30–38 (2013).
13. Borman, G. Examination of a Self-Affirmation Intervention in St. Paul Public Schools. The Council of the Great City Schools, Volume 12 (2012).
14. Dee, T.S. Social Identity and Achievement Gaps: Evidence From an Affirmation Intervention, *Journal of Research on Educational Effectiveness*, 8:2, 149-168 (2015).
15. de Jong, E.M. , Jellesma, F.C., Koomen, H.M.Y.. and de Jong, P.F. A Values-Affirmation Intervention Does Not Benefit Negatively Stereotyped Immigrant Students in the Netherlands. *Frontiers in Psychology*. 7:691 (2016).
16. Bratter, J., Rowley, K.J., and Chukhray, I. Does a Self-Affirmation Intervention Reduce Stereotype Threat in Black and Hispanic High Schools? Race and Social Problems 8, 340-356 (2016).
17. Hoffman, A.J., and Kurtz-Costes, B. Promoting science motivation in American Indian middle school students: an intervention, Educational Psychology, 39:4, 448-469 (2019).

18. Cochran, W.G. Analysis of Covariance: Its Nature and Uses. *Biometrics* 13 (3) 261 (1957).

19. Miller, G.A., Chapman, J.P. Misunderstanding Analysis of Covariance. *J. Abnorm. Psychol.* 110 (1), 40 (2001).

20. Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5), 702-712 (2016).

21. Simonsohn, U., Simmons, J. P., & Nelson, L. D. Specification curve: Descriptive and inferential statistics on all reasonable specifications. Available at SSRN 2694998 (2015).

22. Simmons, P., Nelson, L.D., Simonsohn, U. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci*. 22, 1359-1366 (2011).

23. Open Science Collaboration, An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspect. Psychol. Sci.* 7, 657–660 (2012).

24. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* 349, aac4716 (2015).

25. Simonsohn. Each Reader Decides if a Replication Counts. *Psychol. Sci.* 27 (10) 1410-1412. (2016).

26. Camerer, C.F., Dreber,A., Forsell, E., Ho, T., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., Wu, H. Evaluating replicability of laboratory experiments in economics. *Science* 351, 1433-1436 (2016).