Commentary

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
1–4

DOI: 10.1177/0956797619872762
www.psychologicalscience.org/PS
SAGE

Is There a Gender-Equality Paradox in Science, Technology, Engineering, and Math (STEM)? Commentary on the Study by Stoet and Geary (2018)

Sarah S. Richardson[1,2], Meredith W. Reiches[3], Joe Bruch[4], Marion Boulicault[5], Nicole E. Noll[6], and Heather Shattuck-Heidorn[7]

[1]Department of the History of Science, Harvard University; [2]Committee on Degrees in Studies of Women, Gender, and Sexuality, Harvard University; [3]Department of Anthropology, University of Massachusetts Boston; [4]Department of Population Health Sciences, Harvard University; [5]Department of Philosophy, Massachusetts Institute of Technology; [6]Department of Psychology, Harvard University; and [7]Department of Women and Gender Studies, University of Southern Maine

Received 3/27/19; Revision accepted 7/31/19

In the corrected version of their 2018 article, Stoet and Geary (Corrigendum issued 2019) responded to our identification of a mismatch between their numbers for women in science, technology, engineering, and math (STEM) with tertiary degrees and the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2015) data they sourced. They clarified that their numbers do not represent the percentage of women among STEM graduates, as they had originally stated. Rather, their numbers represent a ratio, which they claim measures the "propensity" for women compared with men to earn a tertiary degree in STEM in a given country (p. 584). The use of this measure in combination with the Global Gender Gap Index (GGGI), a contested (Else-Quest & Hamilton, 2018; Hawken & Munck, 2013) composite measure of nation-level gender equality increasingly employed in similar studies advancing the hypothesis of a gender-equality paradox (e.g., Falk & Hermle, 2018), raises methodological and empirical questions about their claims that there is a gender-equality paradox in STEM and that a larger gender gap in STEM achievement in high gender-equality countries is evidence of baseline sex differences in career and educational preferences.

Propensity to Graduate in STEM

Women's share of STEM degrees relative to men's share of STEM degrees is a logical and interpretable statistic standardly used in scholarship on gender, STEM achievement, and economic development (Ceci, Ginther, Kahn, & Williams, 2014; Ceci, Williams, & Barnett, 2009; Charles & Bradley, 2009; Cheryan, Ziegler, Montoya, & Jiang, 2017; Miller, Eagly, & Linn, 2015) Stoet and Geary instead operationalized nation-level women's participation in STEM as $a/(a + b)$, "where a is the percentage of women who graduate with STEM degrees (relative to all women graduating) and b is the percentage of men who graduate with STEM degrees (relative to all men graduating)" (p. 584).

What does Stoet and Geary's propensity ratio measure? Worldwide, women earn more tertiary degrees than men. In Algeria, 62.7% of tertiary graduates, and 53.55% of STEM graduates, are women. Yet Stoet and Geary reported a value of 40.7 for Algeria. In Poland, 43.63% of STEM graduates are women, which would place it 5th out of the 45 UNESCO countries with Programme for International Student Assessment (PISA) data included in Stoet and Geary's analysis. Yet Stoet and Geary's measure yields a value of 26.9, which ranks Poland 20th in their data set. Conversely, 21.83% of Luxembourg's STEM graduates are women, ranking it 44th among 45 countries, yet Stoet and Geary reported a value of 28.7 for Luxembourg, ranking it 15th.

In their corrected article, Stoet and Geary argued that they intended this measure to capture women's "propensity" to graduate in STEM. We assume that they

Corresponding Author:
Sarah S. Richardson, Harvard University, Department of the History of Science, 1 Oxford St., Science Center 371, Cambridge, MA 02138
E-mail: srichard@fas.harvard.edu

mean propensity in the sense of tendency, inclination, or predisposition, consistent with their hypothesis that men and women have different baseline preferences for STEM study (e.g., Lippa, Collaer, & Peters, 2010; Pinker, 2008; Schmitt, 2015), which can be "exaggerate[d]" and "abated or overridden" (Stoet & Geary, 2018, p. 591) depending on the context. But individual predispositions should be measured by a psychometrically sound scale of individuals' perceived (or actual) gender inequality in relation to their STEM preferences. Both the UNESCO tertiary-degree statistics and GGGI are inappropriate for these purposes because they tell us nothing about individual STEM preferences or individual experiences of gender inequality.

Setting this aside, a propensity variable is conceptually discordant with the GGGI. The GGGI measures achieved outcomes, not propensities (World Economic Forum, 2018, p. 4). The more appropriate measure for analysis of gender equality in STEM tertiary degrees alongside the GGGI is gross completion rates—for example, Algeria's 53.55% or Luxembourg's 21.83%. However, adopting this measure would not fully resolve problems in Stoet and Geary's use of the GGGI, for reasons elaborated below.

Gender Equality

Stoet and Geary characterized the GGGI as a measurement instrument of the degree of gender equality at the nation level, claiming that a high composite index score captures "more educational and empowerment opportunities" for girls and women, including promotion of "girls' and women's engagement in STEM fields" (Stoet & Geary, 2018, p. 591). But the GGGI does not measure opportunity, empowerment, or STEM encouragement. The GGGI is an index for ranking countries by the gap in parity between women and men on select indicators for which there are global, quantitative data available since 2006.

The World Economic Forum (2018) designed GGGI as a "benchmarking tool" to "track a country's progress over time" (p. 32). GGGI provides information about gender equality within countries and ranks countries relative to each other. It is unclear whether a tool designed for ranking, benchmarking, and tracking can be included in correlations without consideration of country-specific parameters and validation of its composite scale as a measure appropriate for correlation with the variable of interest (Else-Quest & Hamilton, 2018; Hawken & Munck, 2013).

By design, the GGGI is not intended to be used to causally explain outcomes, and gender-equal outcomes cannot be interpreted as providing information on causal context within countries. For example, among nations with perfect parity in the health and survival domain are lower-income countries in which both women and men have short life expectancies as well as higher-income countries in which both women and men live longer. Similarly, GGGI rank does not distinguish between top-performing countries in the GGGI political-empowerment subindex, such as France, Ireland, and Slovenia, which have candidate quota laws mandating gender balance, and Germany, Iceland, Norway, which do not (World Economic Forum, 2015; International Institute for Democracy and Electoral Assistance, 2019). In short, the GGGI is neutral with respect to how outcomes of parity are achieved (Thompson, 2017).

Gender equality is a multidimensional construct. Just as a nation's high GGGI score does not mean that gender-equal outcomes are due to gender equality, a low score does not predict gender inequality or gender-unequal outcomes in all domains. Algeria, 128th on the 2015 GGGI (World Economic Forum, 2015), has the world's second-highest rate of STEM tertiary degrees for women. Algeria's extraordinary achievements for women in STEM could result from societal investment in women's education in STEM, distinct cultural beliefs about women's capacity and affinity for STEM, vast overenrollment of women compared with men in tertiary-degree programs, disproportionate outflow of men compared with women to STEM tertiary-degree programs in other Francophone countries, a limited range of tertiary degrees on offer that channel both more women and men into STEM, or other factors (Charles & Bradley, 2009; Thébaud & Charles, 2018).

Results Change Depending on Measures Used

We have argued that the measures of STEM achievement and gender equality that Stoet and Geary used are inappropriate. Does the negative association that they reported persist when measures change?

Method

We hypothesized that the construct of gender equality, the measure of women's achievement in STEM, and countries included in the analysis all affect the correlation between gender equality and women's representation in STEM. As an example, we offer an alternative analysis, using a different measure of gender equality, that proposed by Stoet and Geary (2019), the Basic Index of Gender Inequality (BIGI), to test the robustness of Stoet and Geary's operationalization of gender equality and women's STEM achievement. We changed the measure of women's achievement in STEM to the percentage of women among tertiary STEM graduates as reported by UNESCO, a measure consonant with the focus on achieved outcomes in gender-equality indices.

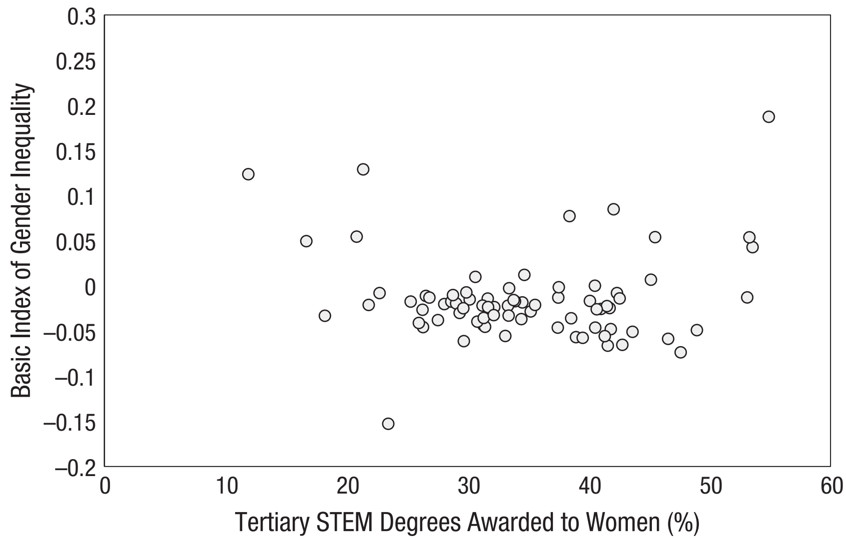

Fig. 1. Scatterplot showing the relationship between the Basic Index of Gender Inequality score (Stoet & Geary, 2019) and the percentage of women among all tertiary graduates in science, technology, engineering, and math (STEM), separately for each of the 77 countries from the United Nations Educational, Scientific and Cultural Organization (UNESCO) 2015 data set.

We included all countries with available BIGI and UNESCO STEM data for the year 2015 (*N* = 77) rather than limiting the analysis to countries with PISA data (*n* = 45). This made possible the inclusion of regions not represented or underrepresented by Stoet and Geary (2018), including sub-Saharan Africa (e.g., Angola, Benin, Burkina Faso), Brazil, and Central and East Asia (e.g., Azerbaijan, Bhutan, Cambodia).

Results

Stoet and Geary (2018) reported a significant correlation between the GGGI and their propensity measure, $r_s = -.47$, 95% CI = [−.66, −.22], $p < .001$, $n = 50$. We conducted a two-tailed bivariate Spearman correlation analysis to assess the relationship between the BIGI and the percentage of women among STEM graduates. The correlation was not significant, $r_s = -.075$, $p = .518$ (Fig. 1). Restricting the analysis to the 45 countries for which BIGI and PISA data are available yielded similar results, $r_s = .070$, $p = .647$. Analysis of the relationship between the BIGI and Stoet and Geary's propensity measure yielded a modest correlation, $r_s = .266$, $p = .021$ ($n = 75$), but this correlation became nonsignificant when restricted to PISA countries, $r_s = .240$, $p = .117$, $n = 44$.

Discussion

The nonsignificant correlation between nation-level gender equality and women in STEM when measures of gender equality and STEM achievement were changed indicates that the association is sensitive to choice of measures of gender equality, range of countries, and STEM achievement. We maintain, however, that these patterns tell us little about global, causal relationships between nation-level measures of gender equality and women in STEM. Substantiating causal relationships between nation-level gender equality and women's STEM achievement would necessitate a longitudinal study design tracking change over time. Further, one would need to account for the nonindependence of countries that share common history, language, and cultures—ideally with a multilevel model that builds in correlations among countries. As argued above, such an analysis would also require a valid measure of experiences of gender inequality in relation to concordant measures of women's STEM achievement and a global range of country-level data.

Transparency

Action Editor: Timothy J. Pleskac
Editor: D. Stephen Lindsay
Author Contributions
S. S. Richardson drafted the manuscript. M. W. Reiches and J. Bruch analyzed the data. All authors provided critical revisions and approved the final version of the manuscript for submission.
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Open Practices

Data used for this analysis were drawn from the same sources used by Stoet and Geary (2018), namely the United Nations Educational, Scientific and Cultural Organization (2015) and World Economic Forum (2015).

ORCID iD

Sarah S. Richardson https://orcid.org/0000-0002-9462-3051

References

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, *15*, 75–141. doi:10.1177/1529100614541236

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, *135*, 218–261. doi:10.1037/a0014412

Charles, M., & Bradley, K. (2009). Indulging our gendered selves? Sex segregation by field of study in 44 countries. *American Journal of Sociology*, *114*, 924–976. doi:10.1086/595942

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*, 1–35. doi:10.1037/bul0000052

Else-Quest, N. M., & Hamilton, V. (2018). Measurement and analysis of nation-level gender equity in the psychology of women. In C. B. Travis & J. W. White (Eds.), *APA handbook of the psychology of women* (Vol. 2, pp. 545–563). Washington, DC: American Psychological Association.

Falk, A., & Hermle, J. (2018). Relationship of gender differences in preferences to economic development and gender equality. *Science*, *362*(6412), Article eaas9899. doi:10.1126/science.aas9899

Hawken, A., & Munck, G. L. (2013). Cross-national indices with gender-differentiated data: What do they measure? How valid are they? *Social Indicators Research*, *111*, 801–838. doi:10.1007/s11205-012-0035-7

International Institute for Democracy and Electoral Assistance. (2019). *Gender quotas database*. Retrieved from https://www.idea.int/data-tools/data/gender-quotas/database

Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior*, *39*, 990–997.

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, *1073*, 631–647. doi:10.1037/edu0000005

Pinker, S. (2008). *The sexual paradox: Men, women and the real gender gap*. New York, NY: Simon & Schuster.

Schmitt, D. P. (2015). The evolution of culturally-variable sex differences: Men and women are not always different, but when they are. . .It appears *not* to result from patriarchy or sex role socialization. In T. K. Shackelford & R. D. Hansen (Eds.), *The evolution of sexuality* (pp. 221–256). London, England: Springer.

Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological Science*, *29*, 581–593. doi:10.1177/0956797617741719

Stoet, G., & Geary, D. C. (2019). A simplified approach to measuring national gender inequality. *PLOS ONE*, *14*(1), Article e0205349. doi:10.1371/journal.pone.0205349

Thébaud, S., & Charles, M. (2018). Segregation, stereotypes, and STEM. *Social Sciences*, *7*(7), Article 111. doi:10.3390/socsci7070111

Thompson, S. (2017, May 2). *How Rwanda beats the United States and France in gender equality*. Retrieved from https://www.weforum.org/agenda/2017/05/how-rwanda-beats-almost-every-other-country-in-gender-equality/

United Nations Educational, Scientific and Cultural Organization. (2015). *Education: Distribution of tertiary graduates by field of study* [Data file]. Retrieved from http://data.uis.unesco.org/index.aspx?queryid=163

World Economic Forum. (2015). *The global gender gap report 2015*. Retrieved from http://www3.weforum.org/docs/GGGR2015/cover.pdf

World Economic Forum. (2018). *The global gender gap report 2018*. Retrieved from http://www3.weforum.org/docs/WEF_GGGR_2018.pdf