Journal of Experimental Social Psychology 61 (2015) 144–147

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

FlashReport

Gender equality probably does not affect performance at the Olympic games: A comment on Berdahl, Uhlmann, and Bai (2015)

Toon Kuppens [a,*], Thomas V. Pollet [b]

[a] Faculty of Behavioural and Social Sciences, University of Groningen, Netherlands
[b] Faculty of Psychology and Education, VU University Amsterdam, Netherlands

HIGHLIGHTS

- Berdahl, Uhlmann, and Bai (2015) analyzed Olympic medals per country.
- They found a positive relation between gender equality and Olympic medals.
- This relation is reduced when controlling for GDP per capita.
- The original analyses also violated the assumption of independence of data points.
- Gender equality is probably not related to performance at the Olympic games.

ARTICLE INFO

Article history:
Received 3 February 2015
Revised 20 April 2015
Accepted 9 June 2015
Available online 10 June 2015

Keywords:
Cross-cultural differences
Statistical analysis
Olympic games
Cross-national comparison
Gender inequality
Galton's problem

ABSTRACT

In a recent article, Berdahl, Uhlmann, and Bai (2015) reported that countries with higher gender equality won more medals at the 2012 and 2014 Olympic games. This relation held for both female and male athletes. The authors, however, did not control for GDP per capita, or take into account the clustering of countries in regions. Here we show that controlling for these two factors reduces or even reverses the positive relation between gender equality and the number of Olympic medals. Gender equality was associated with fewer medals for male athletes. We argue for more careful analyses and interpretation of nation-level data.

1. Introduction

Berdahl, Uhlmann, and Bai (2015), referred to here as BUB, presented data suggesting that countries with greater gender equality win more medals in Olympic games. Importantly, they argue that this relation holds for medals won by women and men. BUB conclude that gender relations are not a zero-sum game, thereby arguing for a causal relation from gender equality to Olympic medals.

Here we would like to comment on BUB's analyses and interpretation thereof, and we highlight two important factors that have not been taken into account by BUB. First, the analyses do not control for GDP per capita. Second, in their analyses countries are assumed to be independent data points, which is not likely to be true and known as Galton's Problem (Kuppens & Pollet, 2014; Mace & Pagel, 1994; Pollet, Tybur, Frankenhuis, & Rickard, 2014). Taking into account these two issues reduces or even reverses the positive relation between gender equality and Olympic medals.

1.1. GDP per capita matters

BUB controlled for GDP and population size separately. This is an unusual choice, given that in cross-cultural research the more logical choice is for GDP per capita, rather than GDP (e.g., Allik & Realo, 2004; Fincher & Thornhill, 2008). This is also the case for research on gender equality (e.g., Inglehart, Norris, & Welzel, 2002). When analyzing Olympic medals, it is essential to control for how wealthy a country is. The wealthier a country is, the more money it can invest per athlete, and this is likely to improve performance. Thus, there is a straightforward theoretical relation between countries' wealth and their performance at the Olympics. Given that the investment per athlete is a measure scaled for population, GDP per capita seems a more logical choice than GDP, especially when population size is already used as a separate

* Corresponding author at: Department of Psychology, Grote Kruisstraat 2/1, 9712TS Groningen, Netherlands.
E-mail address: t.kuppens@rug.nl (T. Kuppens).

http://dx.doi.org/10.1016/j.jesp.2015.06.002
0022-1031/

predictor. We therefore decided to rerun the analyses while controlling for GDP per capita rather than GDP.

1.2. Region matters

As reported in BUB's Table 1, the gender gap score is positively related to women's and men's medals, $r = .22$ and $r = .24$, respectively. However, when the data is split by region, the picture is much less clear. The correlations between gender gap and men's medals range from −.48 in Northern and Western Europe to .30 in Africa (see Table 1). Overall, the weighted average within-region correlation is .006 for women and −.03 for men. At the same time, there are very large differences between regions in the number of male medals ($\eta^2 = .40$), the number of female medals ($\eta^2 = .49$), the gender gap score ($\eta^2 = .51$), and GDP per capita ($\eta^2 = .64$).

This illustrates that region matters for the relation between the gender gap score and Olympic medals. Table 1 also suggests that this relation might be different within regions than between regions. While the within-region relation seems to be close to zero, the between-region correlations between these variables range from .59 (gender gap and female medals) to .76 (men's medals and GDP per capita). Clearly then, relations between these variables reflect between-region dynamics in addition to between-country dynamics, and the existence of regions should therefore be taken into account in the analyses. Table 1 merely illustrates this issue; a more thorough analysis is presented below.

2. Analyses

In our analyses we focus on the overall gender gap score and the education gender gap score (always in two separate models). We chose the education gender gap because it was a stronger predictor of Olympic medals in BUB's analyses, compared to other gender gap dimensions. As in the original BUB paper, we use Poisson regression models, corrected for overdispersion (also referred to as quasipoisson regression models). Correlations between all continuous measures are presented in Table 2.

We replicate the results of BUB although our results are slightly different because we use quasipoisson regression in R and BUB used SAS.[1] For men, the coefficients for the gender gap are .31, $SE = .15$, $p = .04$ for the overall gender gap score and 1.45, $SE = .46$, $p = .002$ for the education gender gap score. For women, we find .44, $SE = .18$, $p = .02$ for the overall gender gap score and 1.63, $SE = .65$, $p = .01$ for the education gender gap score.

2.1. Controlling for GDP per capita

We used the log of GDP per capita in all analyses, as this makes the distribution much less skewed. GDP per capita (standardized) always had a positive relation with the number of medals (all coefficients > .80, all $ps < .003$). When replacing GDP with GDP per capita in the model for men's medals, the coefficient for the overall gender gap score reversed in sign from .31 to −.15 ($SE = .15$, $p = .31$) and the coefficient for the education gender gap score has decreased from 1.44 to .52 ($SE = .46$, $p = .26$).

For women's medals, the results are similar. When adding GDP per capita to the model, the coefficient for the overall gender gap score has decreased from .44 to .06 ($SE = .18$, $p = .74$). The coefficient for the education gap score decreased from 1.63 to .72 ($SE = .73$, $p = .32$).

[1] Using SPSS, we were able to perfectly replicate BUB's results; results for the additional quasipoisson regression models we present here hardly differed between SPSS and R. This means that differences between our and BUB's results are not due to the use of different statistical software.

Table 1
Pearson product–moment correlations between the gender gap score and the number of Olympic medals, by region.

	Women's medals	Men's medals	Number of countries
Africa	.30	.30	24
South and Southeast Asia	−.29	−.32	15
Northern and Western Europe	−.48	−.48	13
South America	.11	.17	12
Central and Eastern Europe	−.30	−.33	12
Central America and Caribbean	.03	.00	11
Central and East Asia	.24	.25	11
Middle East and North Africa		−.06	10
Southern Europe	−.08	.02	9
Anglo-Saxon countries	−.68	−.16	4

2.2. Controlling for region

We used a division in ten world regions and added this to the model as a categorical factor (which is equivalent to adding nine dummy variables). Adding region to the model means that the other variables in the model now assess within-region comparisons rather than a mix of within-region and between-region comparisons. For men's medals, when adding region (10 world regions) to the model, the coefficient for the overall gender gap score has decreased from .31 to −.26 ($SE = .21$, $p = .23$) and the coefficient for the education gender gap score has decreased from 1.44 to .75 ($SE = .44$, $p = .09$).

For women's medals, when adding region (10 world regions) to the model, the coefficient for the overall gender gap score decreased from .44 to −.06 ($SE = .21$, $p = .79$) and the coefficient for the education gender gap score decreased from 1.63 to 1.00 ($SE = .43$, $p = .02$).

2.3. Controlling for both GDP per capita and region

For men's medals, when controlling for both GDP per capita and region (10 world regions), the coefficient for the overall gender gap score has changed from .31 to a negative coefficient, −.55 ($SE = .20$, $p = .007$). The coefficient for the education gender gap score changed from 1.44 to .25 ($SE = .42$, $p = .55$), or less than a quarter of its original size.

For women's medals, when controlling for both GDP per capita and region (10 world regions), the coefficient for the overall gender gap score decreased from .44 to −.26 ($SE = .20$, $p = .19$) and the coefficient for the education gender gap score decreased from 1.63 to .47 ($SE = .39$, $p = .22$).

2.4. Gender gap or GDP per capita?

GDP per capita is strongly related to the gender gap scores (see Table 2), which could have created multicollinearity in the models containing both variables. However, none of the variance inflation factors for the gender gap score or GDP per capita were larger than 2, indicating that multicollinearity was not problematic (Fox, 2008; Fox & Weisberg, 2011).

Another issue raised by the strong relation between GDP per capita and the gender gap score is the conceptual interpretation of their relation with Olympic medals. Does the absence of a direct effect mean that gender equality is irrelevant for Olympic medals, or does gender equality have an indirect effect through GDP per capita? This is similar to asking whether gender equality increases economic development or economic development increases gender equality. The effect probably exists in both directions (Dollar & Gatti, 1999). This could mean that there is a (relatively small) positive indirect effect of gender equality through GDP per capita on Olympic medals. Theoretically, however, such an indirect effect would be incompatible with BUB's proposed explanation of the gender equality effect in terms of gender stereotypes

Table 2
Correlations between continuous measures.

		1	2	3	4	5	6	7	8
1	Women's medals								
2	Men's medals	.82*							
3	Gender gap score	.22**	.24*						
4	Educational gap score	.19**	.23**	.58*					
5	GDP per capita (log)	.37*	.47*	.47*	.61*				
6	GDP	.87*	.64*	.12	.13	.29*			
7	Population	.40*	.28*	−.06	−.05	−.05	.50*		
8	GINI index	−.07	−.20**	−.10	.01	−.21**	.01	.02	
9	Latitude	.19**	.26*	.16	.21**	.40*	.17	.07	−.67*

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

that impede both women's and men's performance. We should also not forget that the direct effect of gender equality is negative for men's medals and that controlling for region also made the gender gap effect much smaller. Overall, the case for an effect of gender gap rather than GDP per capita on male Olympic performance is weak.

2.5. Robustness checks

As a check on the robustness of our findings, we carried out a few additional analyses (details of these analyses can be found in the Supplementary material). We re-ran all models using negative binomial regression rather than quasipoisson regression, we used two different alternative divisions in 12 world regions, and we analyzed data on female and male medals together. All results were similar to the ones presented here.

3. Discussion

The positive relation between gender equality and number of Olympic medals is strongly reduced or reversed after controlling for GDP per capita and region. As such we believe the evidence for a robust statistical association between measures of gender equality and number of Olympic medals is weak at best. In fact, the only reliable relation between gender equality and medals in the quasipoisson models, after controlling for GDP per capita and region, is a negative association between gender gap and men's medals. We do not, however, want to conclude that gender equality impedes men's performance at the Olympic games. In our opinion, information on investment in male and female Olympic athletes would be needed to draw such conclusions, and ideally also individual-level data on motivation and aspirations (e.g., Gill, 1988).

How can problems such as the ones we have highlighted here be avoided? First, when performing nation-level analyses, a lot of attention is needed to select good control variables, as in any correlational study (e.g., McClendon, 1994, pp. 8–10). Although BUB included a range of control variables, they did not include GDP per capita, the most widely used indicator of the economic development of a country. Second, the quality of focal variables must also be checked (e.g., Byrne & Campbell, 1999; Kuppens & Pollet, 2014; Pollet et al., 2014; Poortinga, 1989). The number of Olympic medals can be objectively established, but the gender gap index has some issues. For example, the education gender gap score ranges from .51 to 1 but the median is .99, so there is a strong ceiling effect. Moreover, those who constructed the index have not allowed for the fact that in many countries women outperform men in education to be reflected in the index (see http://reports.weforum.org/global-gender-gap-report-2014/part-1/the-global-gender-gap-index-2014/). Twenty-one countries have a score of 1 on the education gender gap, but most of those countries actually have a score above 1, indicating an educational advantage of women over men. Censoring the measure at 1 seems more likely to be an ideological than a scientific decision. Interestingly, the educational gap was the aspect of the gender gap that predicted the number of medals best. We think that the effects of a variable with such strong skewness should be regarded with caution.

Third, countries are not independent data points, and this needs to be taken into account in the analyses (e.g., Mace & Pagel, 1994; Ross & Homer, 1976). This issue has long been established in the field of cross-cultural research but not in other fields. Countries are clustered in space (and history) and as such they are like members of the same family or pupils of the same classroom: they tend to be similar to each other. This similarity violates the statistical assumption of independence. Many solutions to this problem exist (e.g., spatial modeling via GIS, Chang, 2003) but a simple and first robustness check can consist of controlling for region (e.g., Kuppens & Pollet, 2014), or analyzing the data at a higher level (e.g., region). If such an analysis does not uphold the statistical patterns, then geographical or cultural clustering could drive the found effect and it is less likely that it reflects a real phenomenon.

We have focused here on GDP per capita and the non-independence of countries. There might also be other issues that affect the interpretation of cross-national data such as these. For example, one can question whether a science that focuses on psychological processes can learn much from an exclusive focus on nation-level analyses. The correspondence between those levels, individual and national, can be low. Oishi and Diener (2014), for example, had a hypothesis and data at the individual level, but they surprisingly tested their hypothesis at the national level (see Kuppens & Pollet, 2014). While Oishi and Diener (2014) reported a negative relationship between wealth and meaning of life at the country level, the relationship at the individual level is actually positive. Regarding Berdahl et al. (2015), even if there was a robust, nation-level statistical association between gender equality and Olympic performance, this does not imply anything about the effect of gender equality on individual athletes. The potential lack of correspondence between analyses at some aggregate level and the individual level is known as the "ecological fallacy" (e.g., Connolly, 2006; Pearce, 2000; Robinson, 1950; Yip & Liu, 2006).

In conclusion, we call for more careful assessment of relationships at country level, such as in Berdahl et al. (2015) and what they could entail for individual psychology. Country level data clearly differs from individual, psychological data in many respects and we call for caution when interpreting and inferring patterns from such data.

Acknowledgments

The authors would like to thank Jennifer Berdahl, Eric Uhlmann, and Feng Bai for their comments on an earlier version of this article, and Mike Friedman for the help with the statistical analyses. Toon Kuppens is supported by a University of Groningen endowed chair for Prof. Russell Spears [170240105].

Appendix A. Supplementary material

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jesp.2015.06.002.

References

Allik, J., & Realo, A. (2004). Individualism–collectivism and social capital. *Journal of Cross-Cultural Psychology*, *35*(1), 29–49. http://dx.doi.org/10.1177/0022022103260381.

Berdahl, J.L., Uhlmann, E.L., & Bai, F. (2015). Win–win: Female and male athletes from more gender equal nations perform better in international sports competitions. *Journal of Experimental Social Psychology*, *56*, 1–3. http://dx.doi.org/10.1016/j.jesp.2014.08.003.

Byrne, B.M., & Campbell, T.L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure. *Journal of Cross-Cultural Psychology*, *30*(5), 555–574. http://dx.doi.org/10.1177/0022022199030005001.

Chang, K. -T. (2003). *Introduction to geographic information systems*. New York, NY: McGraw-Hill.

Connolly, P. (2006). Summary statistics, educational achievement gaps and the ecological fallacy. *Oxford Review of Education*, *32*(02), 235–252.

Dollar, D., & Gatti, R. (1999). *Gender inequality, income, and growth: Are good times good for women? Policy research report on gender and development working paper series; no. 1*Washington, DC: The World Bank (Retrieved from http://documents.worldbank.org/curated/en/1999/05/443556/gender-inequality-income-growth-good-times-good-women).

Fincher, C.L., & Thornhill, R. (2008). Assortative sociality, limited dispersal, infectious disease and the genesis of the global pattern of religion diversity. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1651), 2587–2594. http://dx.doi.org/10.1098/rspb.2008.0688.

Fox, J. (2008). *Applied regression analysis and generalized linear models* (Second ed.). Los Angeles, CA: Sage.

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*. Thousand Oaks, CA: Sage.

Gill, D.L. (1988). Gender differences in competitive orientation and sport participation. *International Journal of Sport Psychology*, *19*(2), 145–159.

Inglehart, R., Norris, P., & Welzel, C. (2002). Gender equality and democracy. *Comparative Sociology*, *1*(3), 321–345. http://dx.doi.org/10.1163/156913302100418628.

Kuppens, T., & Pollet, T.V. (2014). Mind the level: Problems with two recent nation-level analyses in psychology. *Frontiers in Psychology*, *5*, 1110. http://dx.doi.org/10.3389/fpsyg.2014.01110.

Mace, R., & Pagel, M. (1994). The comparative method in anthropology. *Current Anthropology*, *35*(5), 549–564.

McClendon, M.J. (1994). *Multiple regression and causal analysis*. Itasca, IL: Peacock.

Oishi, S., & Diener, E. (2014). Residents of poor nations have a greater sense of meaning in life than residents of wealthy nations. *Psychological Science*, *25*, 422–430. http://dx.doi.org/10.1177/0956797613507286.

Pearce, N. (2000). The ecological fallacy strikes back. *Journal of Epidemiology and Community Health*, *54*(5), 326–327.

Pollet, T.V., Tybur, J.M., Frankenhuis, W.E., & Rickard, I.J. (2014). What can cross-cultural correlations teach us about human nature? *Human Nature (Hawthorne, N.Y.)*, *25*(3), 410–429. http://dx.doi.org/10.1007/s12110-014-9206-3.

Poortinga, Y.H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology*, *24*(1), 737–756. http://dx.doi.org/10.1080/00207598908247842.

Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*(3), 351–357.

Ross, M.H., & Homer, E. (1976). Galton's problem in cross-national research. *World Politics*, *29*(1), 1–28. http://dx.doi.org/10.2307/2010045.

Yip, P.S.F., & Liu, K.Y. (2006). The ecological fallacy and the gender ratio of suicide in China. *The British Journal of Psychiatry*, *189*(5), 465–466. http://dx.doi.org/10.1192/bjp.bp.106.021816.