LETTER

Dutch research funding, gender bias, and Simpson's paradox

Based on, among other criteria, three consecutive years of grant applications to the "Veni programme" of the Netherlands Organization for Scientific Research (NWO), van der Lee and Ellemers (1) conclude that these data "provide compelling evidence of gender bias in personal grant applications to obtain research funding." One of the main results this claim is based upon is that of the 1,635 applications by males, 17.7% were successful, whereas of the 1,188 applications by females, only 14.4% were successful. When applying the χ^2 test to the data, the authors found a P value of 0.045 (1). This conclusion is based on the application of an inappropriate statistical procedure, and is therefore questionable, due to the so-called "Simpson's paradox."

Simpson's paradox dates back to Simpson (2) and gained familiarity after a report by Bickel et al. (3). In essence, Simpson's paradox states that an apparently significant relation between two variables in a contingency table can be due to a joint dependency on a third variable.

In table S1 of the article by van der Lee and Ellemers (1), a breakdown of grant applications per research discipline is presented. The proportion of female applicants varies from 11.8% (physics) to 51.4% (health sciences), and the total success rate varies from 13.4% (social sciences) to 26.3% (chemical sciences). Fig. 1 visualizes these data and immediately shows a clear negative relation between the proportion of female applicants and the total success rate. It turns out that women tend to apply more often to competitive disciplines, such as health sciences and social sciences, whereas men apply more often to less competitive disciplines, such as physics and chemical sciences. In four of the nine disciplines, women have a higher success rate than men, and in five of the nine disciplines, men have a higher success rate than women. When taking into account that multiple comparisons are performed, the gender bias is significant (at the $\alpha = 0.05$ level) for none of the disciplines. Thus, when taking into account the spurious correlation, the statistical significance of the relation vanishes.

As a consequence, the conclusion of van der Lee and Ellemers (1) could be severely influenced by this issue, at least for the (large) part where this claim is based on the supposed P value of 0.045. The authors perform other tests on various aspects of the proposal review process, but the data provided in the paper are insufficient to check whether these tests suffer from Simpson's paradox as well. Because these tests study related issues, it is not unlikely that they indeed suffer from this paradox.

Further scrutiny of NWO data are needed to assess whether gender bias indeed affects Dutch science funding. This scrutiny includes studying data from other years and programs. Interestingly, on the aggregate level, men received relatively more grants than women in eight of the 14 Veni calls, with the difference, again, not being statistically significant (4).

ACKNOWLEDGMENTS. This contribution has benefitted from communications with Daniël Lakens, Jelte Wicherts, and Ivo Krausz.

Casper J. Albers[1]

Heymans Institute for Psychological Research, University of Groningen, 9712 TS Groningen, The Netherlands

1 van der Lee R, Ellemers N (2015) Gender contributes to personal research funding success in The Netherlands. *Proc Natl Acad Sci USA* 112(40):12349–12353.

2 Simpson EH (1951) The interpretation of interaction in contingency tables. *J R Stat Soc, B* 13:238–241.

3 Bickel PJ, Hammel EA, O'Connell JWO (1975) Sex bias in graduate admissions: Data from berkeley. *Science* 187(4175): 398–404.

4 Albers CJ (2015) NWO, Gender Bias and Simpson's Paradox. Available at blog.casperalbers.nl/science/nwo-gender-bias-and-simpsons-paradox/. Accessed November 23, 2015.

Author contributions: C.J.A. analyzed data and wrote the paper.

The author declares no conflict of interest.

[1]Email: c.j.albers@rug.nl.

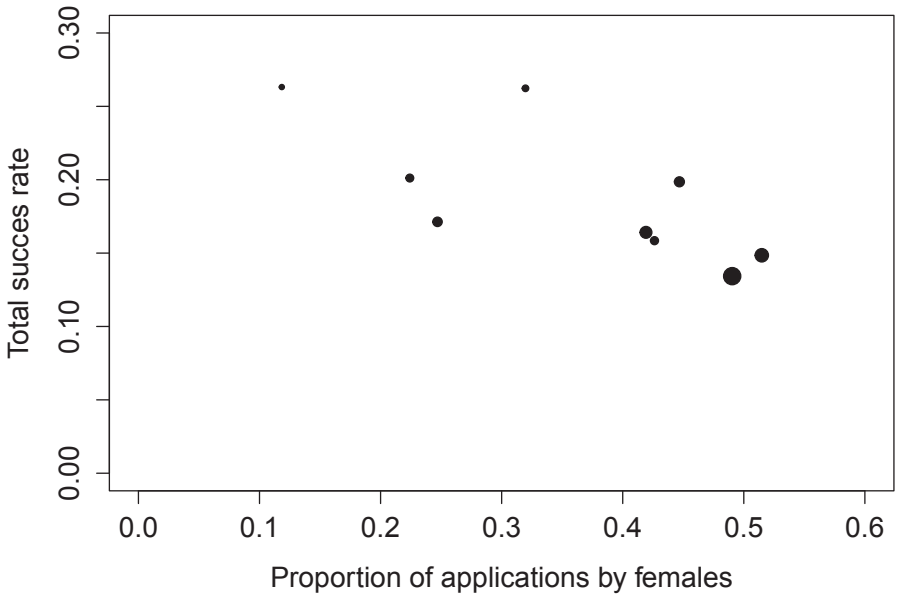

Fig. 1. Proportion of applications submitted by females vs. the (overall) proportion of accepted proposals, per NWO discipline. The size of the markers is proportional to the number of applications.