Research Policy xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Research Policy

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

Short communication

# The big consequences of small biases: A simulation of peer review

Theodore Eugene Day*

*The Children's Hospital of Philadelphia, Office of Safety and Medical Operations, 3401 Civic Center Blvd, Philadelphia, PA 19104, USA*

ARTICLE INFO

*Article history:*
Received 11 August 2014
Received in revised form 12 January 2015
Accepted 16 January 2015
Available online xxx

*Keywords:*
Bias
Discrete event simulation
Grant funding
Peer review

ABSTRACT

*Objective:* To determine the effect of reviewer bias on grant application funding rates between a "preferred class" (PC) and a "non-preferred class" (NPC) of principal investigator.
*Methods:* A discrete event simulation (DES) of grant review was developed which mimics the production, review, and funding determination of grants. Grants were defined to have an intrinsic quality. Three reviewers then score each grant, and assign it a value. Zero (control), one, or all reviewers may exhibit biases of varying severity against NPC investigators.
*Results:* When total review bias exceeds 1.9% of grant score, statistically significant variation in scores between PC and NPC investigators is discernable in a pool of 2000 grant applications. When total review bias exceeds 2.8% of total grant score, statistically significant discrepancies in funding rates between PC and NPC investigators are detectable in a simulation of grant review.
*Conclusions:* Review bias affects funding rates even when total review bias is less than half the amplitude of normal variation in an individual reviewer's score. Addressing reviewer bias will improve equity among investigators and may improve the overall quality of funded grant applications.


## 1. Introduction

Bias in grant application review and award has been the subject of considerable attention, with a recent study of R01 grants awarded by the National Institutes of Health (NIH) showing that Asian and African–American investigators were significantly less likely to receive funding (Ginther et al., 2011) than white counterparts. Another study from 2010 data shows that while overall NIH success rates are similar for male and female researchers, male investigators are significantly more likely to receive follow-up funding when applying as previous NIH grantees (Pohlhaus et al., 2011). Meta-analysis has similarly shown that statistically significant gender bias, favoring males, exists in the peer review of grant applications (Bornmann et al., 2007). Double-blind review has been shown to increase publication rates of articles with female first authors (Budden et al., 2008), and unblinded review has been shown to favor researchers from the United States (Ross et al., 2006). Discipline specific analyses of funding in otolaryngology (Eloy et al., 2013) and physiology (Gordon, 2014) have come to similar conclusions regarding the presence of gender bias.

Strides in combating bias have certainly been made in some settings. Brooks and Della Salla report in the journal Cortex that representative parity had been achieved by that journal as of 2009 (Brook and Della Salla, 2009). Nature Neuroscience reported in 2006 that in a sample of papers submitted to that journal the prior year, acceptance rates among male and female first authors were statistically indistinguishable (Editorial, 2006). These results are promising, and suggest that identifying, studying, and understanding bias can lead to the amelioration of its effects.

Thus, it is known that bias exists, and has been reported to affect both publication and grant application funding. It has been posited that peer-review success is related to "accumulated advantages" related to ethnicity, gender, and professional standing (Vinera et al., 2004). Indeed overt and covert bias continue to plague science, medicine, and academia generally (Rosser, 2013; Cochran et al., 2014). In this investigation, the author attempts to determine at what level quantifiable bias results in statistically significant discrepancies in funding rates between preferred and non-preferred investigators in a discrete event simulation which mimics the peer review of grant applications.

## 2. Methods

This study was a simulated prospective controlled trial of a grant review process. The control and the experimental arms differ only in the presence of bias in review. The design and structure of the simulation are:

2000 grant applications per simulation run were generated in groups of 100, 50 from "preferred class" (PC) investigators and 50 from "non-preferred class" (NPC) investigators. Each grant is

* Tel.: +1 2155900879.
*E-mail address:* dayt@email.chop.edu

http://dx.doi.org/10.1016/j.respol.2015.01.006
0048-7333/© 2015 Elsevier B.V. All rights reserved.

**Table 1**
Results of biased and unbiased review of PC and NPC grant applications.

| | Investigator class | Bias | Intrinsic quality | *P* | Score | *P* | N Funded | With 10% funding line<br>Odds ratio | 95% CI | χ2 | N funded | With 20% funding line<br>Odds ratio | 95% CI | χ2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Control run | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 99 | | | | 196 | | | |
| | Non preferred | 0 | 5.02 ± 1.53 | 0.82 | 5.05 ± 1.85 | 0.86 | 101 | 1.022 | 0.763–1.370 | – | 204 | 1.051 | 0.844–1.309 | – |
| Biased runs | | | | | | | | | | | | | | |
| 1 | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 104 | | | | 200 | | | |
| | Non preferred | 1 @ 0.1 | 5.02 ± 1.53 | 0.82 | 5.08 ± 1.85 | 0.57 | 96 | 0.915 | 0.683–1.226 | 0.57 | 200 | 1.000 | 0.803–1.245 | 1 |
| 2 | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 106 | | | | 202 | | | |
| | Non preferred | 1 @ 0.25 | 5.02 ± 1.53 | 0.82 | 5.13 ± 1.85 | 0.24 | 94 | 0.875 | 0.653–1.173 | 0.4 | 198 | 0.975 | 0.783–1.214 | 0.82 |
| 3 | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 109 | | | | 209 | | | |
| | Non preferred | 1 @ 0.5 | 5.02 ± 1.53 | 0.82 | 5.22 ± 1.85 | **<0.05** | 91 | 0.818 | 0.610–1.097 | 0.2 | 191 | 0.894 | 0.718–1.112 | 0.31 |
| 4 | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 113 | | | | 216 | | | |
| | Non preferred | 1 @ 0.75 | 5.02 ± 1.53 | 0.82 | 5.30 ± 1.85 | **<0.01** | 87 | 0.748 | 0.557–1.004 | 0.07 | 184 | 0.818 | 0.657–1.020 | 0.07 |
| 5 | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 118 | | | | 222 | | | |
| | Non preferred | 1 @ 1 | 5.02 ± 1.53 | 0.82 | 5.38 ± 1.85 | **<0.0001** | 82 | **0.668** | **0.496–0.898** | **0.01** | 178 | **0.759** | **0.609–0.946** | **<0.05** |
| 6 | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 117 | | | | 219 | | | |
| | Non preferred | 3 @ 0.3 | 5.02 ± 1.53 | 0.82 | 5.35 ± 1.83 | **<0.001** | 83 | **0.683** | **0.508–0.918** | **0.02** | 181 | **0.788** | **0.633–0.982** | **<0.05** |
| 7 | Preferred | | 5.03 ± 1.55 | | 5.04 ± 1.83 | | 141 | | | | 255 | | | |
| | Non preferred | 3 @ 0.75 | 5.02 ± 1.53 | 0.82 | 5.8 ± 1.85 | **<0.0001** | 59 | **0.382** | **0.278–0.525** | **<0.0001** | 145 | **0.495** | **0.395–0.621** | **<0.0001** |