A proxy for studying positive-outcome bias without meta-analysis

Daniele Fanelli¹

¹<u>dfanelli@staffmail.ed.ac.uk</u>

The University of Edinburgh, ISSTI, Old Surgeons' Hall, EH1 1LZ, Edinburgh, UK

Introduction

Publication and related biases in the literature scientific are increasingly documented, raising concerns for the reliability of scientific knowledge. Such biases are mostly detected in formal metaanalyses, and this limits our ability to understand their nature and causes. Flexible tools, able to assess bias in random samples of papers, would allow a more systematic study of the phenomenon. Here I overview some of the results obtained with a simple proxy of positiveoutcome bias.

Methods

"test* Sampling: the sentence the hypothes*" was used to search all 10,837 journals available in the Essential Science which Indicators database. classifies journals univocally in 22 disciplines. The mathematics, discipline of however. vielded no usable paper, while the "multidisciplinary" category, which includes journals like Science and Nature, was excluded. Therefore, 20 disciplines were included in the analysis. Up to 150 papers were sampled at random from each discipline.

Coding: by examining the abstract and/or full-text, the specific hypothesis tested in each paper was identified, and it was determined whether the authors had concluded to have found a positive (full or partial) or negative (null or negative) support. If more than one hypothesis was being tested, only the first one listed in the text was considered. Meeting abstracts, papers that did not actually test a hypothesis (around 12% of the total), and those for which information was

insufficient to determine the outcome (around 1%) were excluded.

Reliability: All data was extracted by the author, but an untrained assistant who was given basic written instructions scored papers the same way as the author in 18 out of 20 cases, and picked up exactly the same sentences for hypothesis and conclusions in all but three cases. The discrepancies were easily explained, showing that the procedure is objective and replicable.

Paper characteristics: Any information except discipline (i.e. corresponding address, year of publication etc...) could be retrieved after all papers had been coded, allowing the coding to completely blind to all other paper characteristics.

Statistical analysis: The ability of independent variables to predict the outcome of a paper was tested by standard logistic regression analysis, fitting a model in the form:

logit (Y)=ln $\left[\frac{\beta}{2}\right]$ = $\beta_0 + \beta_1 X_{i1} + ... + \beta_n X_{in}$

in which p_i is the probability of the *i*th paper to report a positive result, and X_1-X_n represent relevant characteristics of the *i*th paper.

Results

The number of papers retrieved from each discipline varied considerably (Table 1). The number of papers retrieved from each discipline increased with number of journals available and was higher in the life sciences (multiple regression: $b=0.59\pm0.24$, P=0.026, $b=2.29\pm0.42$, P<0.001, respectively).

Table 1 reports numbers for the period 1990-1999, because information for the period 2000-2007 was incomplete. The numbers for 2000-2007 were on average

higher, and their relative distribution among disciplines very similar.

Table 1. Papers that in the abstract contained the sentence "test* the hypoth*", published in 1990-1999 in journals from the ESI database.

Discipline	Ν	Discipline	Ν
Agricult Sc	202	Molec Biol	891
Biol&Bioch	2300	Microbiology	229
Chemistry	88	Materials Sc	26
Clinical Med	14181	Neurosc&Beh	1987
Computer Sc	38	Plant&An Sc	1701
Econ&Busin	359	Physics	92
Env&Ecol	954	Psycol&Psychi	1199
Engineering	101	Phar&Toxicol	574
Geosciences	213	Social Sc,Gen	596
Immunology	463	Space Science	79

By sampling each discipline (N \leq 150), and determining each paper outcome, interesting patterns were detected. The frequency of positive results increased with the putative "softness" of scientific disciplines and methodologies, particularly among pure disciplines, whilst applied disciplines showed no pattern (Figures 1 and 2).

The frequency of positive results varied between countries, with the US having fewer positive results than Asian countries, but significantly more than the rest of the world (Figure 3).

Moreover, within the US states, the frequency of positive result was found to be higher in states that, according to NSF data, publish more papers with same or less research funding (Figure 4).

Recent analyses show that the frequency of positive results has been growing over the years, with significant differences between disciplines and countries (Fanelli, unpublished).

Conclusions

The link with factors predicted to increase scientific bias (i.e. the "softness" of research and the productivity of researchers) suggests that this is a genuine proxy of scientific bias, able to reveal patterns that cannot be explained by pure "file-drawer" effects (i.e. differences observed between US and European countries).

Therefore, the proxy can be used to study patterns and causes of positive-outcome bias, in virtually any discipline or country. The small number of papers retrieved in some disciplines in ESI (Table 1) represents a limitation, which might be partly avoided by using other databases.



70% Physical Sciences Biological Sciences Social Sciences (tot N=642) (tot N=1302) (tot N=402)

Figure 1. Percentage of papers that supported the tested hypothesis, by scientific domain. Papers were published in 2000-2007. Bars are logit-derived 95% CI. Odds ratios (95% CI) are for pure and applied combined, corrected for various confounders. Lines were added to visualise trends. Modified from Fanelli (2010a).



Figure 2) Percentage of papers that supported the tested hypothesis, by methodology. Papers were published in 2000-2007. Bars are logit-derived 95% CI. Only pure disciplines are shown, but odds ratios (95% CI) are for pure and applied disciplines combined, corrected for various confounders. Lines were added to visualise trends. Modified from Fanelli (2010a).

Fanelli



Figure 3) Number of papers that supported (white, with percentage) or did not support the tested hypothesis, by geographical location based on corresponding address. Papers were published in 2000-2007. Odds ratios (95% CI) are uncorrected for confounders. "Positive" Results by Academic Productivity



Figure 4) Percentage of papers that supported the tested hypothesis, by productivity of US state (data from the National Science Foundation). States are indicated by USPS abbreviations. Papers were published in 2000-2007. Regression estimates are from multiple logistic regression, corrected for various confounders, including academic R&D expenditure. Modified from Fanelli (2010b).

References

- Fanelli, D. (2010a). "Positive" Results Increase Down the Hierarchy of the Sciences. *Plos One*, 5(3).
- Fanelli, D. (2010b). Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *Plos One*, 5(4).