

No evidence of publication bias in climate change science

Christian Harlos¹ · Tim C. Edgell² · Johan Hollander¹ 

Received: 17 May 2016 / Accepted: 11 December 2016 / Published online: 28 December 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Non-significant results are less likely to be reported by authors and, when submitted for peer review, are less likely to be published by journal editors. This phenomenon, known collectively as publication bias, is seen in a variety of scientific disciplines and can erode public trust in the scientific method and the validity of scientific theories. Public trust in science is especially important for fields like climate change science, where scientific consensus can influence state policies on a global scale, including strategies for industrial and agricultural management and development. Here, we used meta-analysis to test for biases in the statistical results of climate change articles, including 1154 experimental results from a sample of 120 articles. Funnel plots revealed no evidence of publication bias given no pattern of non-significant results being under-reported, even at low sample sizes. However, we discovered three other types of systematic bias relating to writing style, the relative prestige of journals, and the apparent rise in popularity of this field: First, the magnitude of statistical effects was significantly larger in the abstract than the main body of articles. Second, the difference in effect sizes in abstracts versus main body of articles was especially pronounced in journals with high impact factors. Finally, the number of published articles about climate change and the magnitude of effect sizes therein both increased within 2 years of the seminal report by the Intergovernmental Panel on Climate Change 2007.

1 Introduction

Publication bias in scientific journals is widespread (Fanelli 2012). It leads to an incomplete view of scientific inquiry and results and presents an obstacle for evidence-based

Data archiving Data used in the meta-analysis are archived in the Dryad repository

Tim C. Edgell and Johan Hollander contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10584-016-1880-1) contains supplementary material, which is available to authorized users.

✉ Johan Hollander
Johan.Hollander@biol.lu.se

¹ Department of Biology, Aquatic Ecology, Lund University, 223 62 Lund, Sweden

² Stantec, 11-2042 Mills Road, Sidney, BC V8L 5X4, Canada

decision-making and public acceptance of valid, scientific discoveries and theories. A growing trend in scientific inquiry, as practiced in this article, includes the meta-analysis of large bodies of literature, a practice that is particularly susceptible to misleading and inaccurate results given a systematic bias in the literature (e.g. Michaels 2008; Fanelli 2012, 2013).

The role of publication bias in scientific consensus has been described in a variety of scientific disciplines, including but not limited to medicine (Kicinski 2013; Kicinski et al. 2015), social science (Fanelli 2012), ecology (Palmer 1999), and global climate change research (Michaels 2008; Reckova and Irsova 2015).

Despite widespread consensus among climate scientists that global warming is real and has anthropogenic roots (e.g., Holland 2007; Idso and Singer 2009; Anderegg et al. 2010), several end users of science such as popular media, politicians, industrialists, and citizen scientists continue to treat the facts of climate change as fodder for debate and denial. For example, Carlsson-Kanyama and Hörnsten Friberg (2012) found only 30% of politicians and directors from 63 Swedish municipalities believed humans contribute to global warming; 61% of respondents were uncertain about the causes of warming, and as much as 9% denied it was real.

Much of this skepticism stems from an event that has been termed *Climategate*, when emails and files from the Climate Research Unit (CRU) at the University of East Anglia were copied and later exposed for public scrutiny and interpretation. Climate change skeptics claimed the IPCC 2007 report—the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC 2007), which uses scientific facts to argue humans are causing climate change—was based on an alleged bias for positive results by editors and peer reviewers of scientific journals; editors and scientists were accused of suppressing research that did not support the paradigm for carbon dioxide-induced global warming. In 2010, the CRU was cleared by the Muir Russell Committee of any scientific misconduct or dishonesty (Adams 2010; but see Michaels 2010).

Although numerous reviews have examined the credibility of climate researchers (Anderegg et al. 2010), the scientific consensus on climate change (Doran and Kendall Zimmerman 2009) and the complexity of media reporting (Corner et al. 2012), few studies have undertaken an empirical review of the publication record to evaluate the existence of publication biases in climate change science. However, Michaels (2008) scrutinized the two most prestigious journals, *Nature* and *Science*, in the field of global warming, and by using vote-counting meta-analysis, confirmed a skewed publication record. Reckova and Irsova (2015) also detected a publication bias after analyzing 16 studies of carbon dioxide concentrations in the atmosphere and changes in global temperature. Although publication biases were reported by Michaels (2008) and Reckova and Irsova (2015), the former test used a small set of pre-defined journals to test the prediction, while the latter test lacked statistical power given a sample size of 16 studies. In contrast, here we conducted a meta-analysis on results from 120 reports and 31 scientific journals. Our approach expands upon the conventional definition of publication bias to include publication trends over time and in relation to seminal events in the climate change community, stylistic choices made by authors who may selectively report some results in abstracts and others in the main body of articles (Fanelli 2012) and patterns of effect size and reporting style in journals representing a broad cross-section of impact factors.

We tested the hypothesis of bias in climate change publications stemming from the under-reporting of non-significant results (Rosenthal 1979) using fail-safe sample sizes, funnel plots, and diagnostic patterns of variability in effect sizes (Begg and Mazumdar 1994; Palmer 1999, 2000; Rosenberg 2005). More specifically, we (a) examined whether non-significant

results were omitted disproportionately in the climate change literature, (b) if there were particular trends of unexpected and abrupt changes in the number of published studies and reported effects in relation to IPCC 2007 and Climategate, (c) whether effects presented in the abstracts were significantly larger than those reported in the main body of reports, and (d) how findings from these first three tests related to the impact factor of journals.

Meta-analysis is a powerful statistical tool used to synthesize statistical results from numerous studies and to facilitate general trends in a field of research. Unfortunately, not all articles within a given field of science will contain statistical estimates required for meta-analysis (e.g., estimate of effect size, error, sample size). Therefore, the literature used in meta-analysis is often a sample of all available articles, which is analogous to the analytical framework used in ecology and typically uses a sub-sample of a population to estimate parameters of true populations. For the purpose of our meta-analysis, we sampled articles from the body of literature that explores the effects of climate change on marine organisms. Marine species are exposed to a large array of abiotic factors that are linked directly to atmospheric climate change. For instance, oceans absorb heat from the atmosphere and mix with freshwater run-off from melting glaciers and ice caps, which changes ocean chemistry and puts stress on ocean ecosystems. For example, the resulting changes in ocean salinity and pH can inhibit calcification in shell-bearing organisms that are either habitat-forming (e.g., coral reefs, oyster reefs) or the foundation of food webs (e.g., plankton) (The Copenhagen Diagnosis 2009).

Results of our meta-analysis found no evidence of publication bias, in contrast to prior studies that were based on smaller sample sizes than used here (e.g., Michaels 2008; Reckova and Irsova 2015). We did, however, discover some interesting patterns in the numbers of climate change articles being published over time and, within journal articles, stylistic biases by authors with respect to reporting large statistically significant effects. Finally, results are discussed in the context of social responsibility borne by climate scientists and the challenges for communicating science to stakeholders and end users.

2 Materials and methods

Meta-analysis is a suite of data analysis tools that allow for quantitative synthesis of results from numerous scientific studies, now widely used from medicine to ecology (Adams et al. 1997). Here, we randomly sampled articles from a broader body of literature about climate change in marine systems and withdrew statistics summarizing magnitude of effects, error, and experimental sample size for meta-analysis.

2.1 Data collection

We surveyed the scientific literature via the ISI Web of Science, Scopus and Biological Abstracts, and in the reference sections of identified articles for experimental results pertaining to climate change in ocean ecosystems. The search was performed with no restrictions on publication year, using different combinations of the terms: (acidification* AND ocean*) OR (acidification* AND marine*) OR (global warming* AND marine*) OR (global warming* AND ocean*) OR (climate change* AND marine* AND experiment*) OR (climate change* AND ocean* AND experiment*). The search was performed exclusively on scientific journals with an impact factor of at least 3 (Journal Citation Reports science edition 2010).

We restricted our analysis to a sample of articles reporting (a) an empirical effect size between experimental and control group, (b) a measure of statistical error, and (c) a sample size with specific control and experimental groups (see Supplementary Material S1–S3 for identification of studies). We identified 120 articles from 31 scientific journals published between the years 1997 and 2013, with impact factors ranging from 3.04 to 36.104 and a mean of 6.58. Experimental results ($n = 1154$) were extracted from the main body of articles; 362 results were also retrieved from the articles' abstracts or summary paragraphs.

Data from the main body of articles and abstracts were analyzed separately to test for potential stylistic biases related to how authors report key findings. The two datasets, hereafter designated "main" dataset and "abstract" dataset, were also divided into three time periods: pre-IPCC 2007 (x–2007–November), After IPCC 2007/pre-Climategate (2007–December–2009–November) and after Climategate (2009–December–December 2012), based on each article's date of acceptance. We used November 2007 as the publication date for the IPCC Fourth Assessment Report, which was an updated version of the original February 2007 release.

We extracted graphical data using the software GraphClick v. 3.0 (2011). Each study could include several experimental results, which could result in non-independence bias driven by studies with relatively large numbers of results. Therefore, we assessed the robustness of our meta-analysis by re-running the analysis multiple times with data subsets consisting of one randomly selected result per article (Hollander 2008).

Experimental results found in articles can be either negative or positive. To prevent composite, mean values from equalling zero, we reversed the negative sign of effects to positive; consequently, all results were analyzed as positive effects (Hollander 2008). The reversed effect sizes do not generally produce a standard normal distribution, as negative effects are reversed around zero. Statistical significance was therefore assessed using bias-corrected 95% bootstrap confidence intervals produced by re-sampling tests in 9999 iterations, with a two-tailed critical value from Students t distribution. If the mean of one sample lies outside the 95% confidence intervals of another mean, the null hypothesis that subcategories did not differ was rejected (Adams et al. 1997; Borenstein et al. 2010). Hedges' d was used to quantify the weighted effect size of climate change effects (Gurevitch and Hedges 1993).

$$d = \frac{\bar{X}^E - \bar{X}^C}{s} J$$

Hedges' d was the mean of the control group (\bar{X}^C) subtracted from the mean of the experimental group (\bar{X}^E), divided by the pooled standard deviation (s) and multiplied by a correction factor for small sample sizes (J). However, since sample sizes vary among studies, and variance is a function of sample size, some form of weighting was necessary. In other words, studies with larger sample sizes are expected to have lower variances and will accordingly provide more precise estimates of the true population effect size (Hedges and Olkin 1985; Shadish and Haddock 1994; Cooper 1998). Therefore, a weighted average was used in the meta-analysis to estimate the cumulative effect size (weighted mean) for the sample of studies (see Rosenberg et al. 2000 for details).

2.2 Funnel plot and fail-safe sample sizes

Funnel plots are sensitive to heterogeneity, which is why they are effective for visual detection of systematic heterogeneity in the publication record. For example, funnel plot asymmetry has long

been equated with publication bias (Light and Pillemer 1984; Begg and Mazumdar 1994), whereas a systematic inverted funnel is diagnostic of a “well-behaved” data set in which publication bias is unlikely. Initial funnel plots showed large amounts of variation in the y -axis (Hedge’s d) along the length of the x -axis (sample size), which could potentially obscure inspection of diagnostic features of the funnel around the mean. To improve detectability of publication bias, should one truly exist, we transformed Hedges’ d to Pearson coefficient correlation (r), which condensed extreme values in the y -axis and converted the measure of effect size to a range between zero and ± 1 (Palmer 2000; Borenstein et al. 2009). Therefore, the data transformation ultimately converted the measure of effect size from a standardized mean difference (d) to a correlation (r)

$$r = \frac{d}{\sqrt{d^2 + a}}$$

where a is the correction factor for cases where $n_1 \neq n_2$ (Borenstein et al. 2009).

Both funnel plots and fail-safe sample size were inspected to test for under-reporting of non-significant effects, following Palmer (2000). Extreme publication bias (caused by under-reporting of non-significant results) would appear as a hole or data gap in a funnel plot. Also, if there is no bias, the density of points should be greatest around the mean value and normally distributed around the mean at all sample sizes. To help visualize the threshold between significant and non-significant studies, 95% significance lines were calculated for the funnel plots.

Robust Z -scores were used to identify possible outliers in the dataset, as such values could distort the mean and make the conclusions of a study less accurate or even incorrect. Instead of using the dataset mean, robust Z -scores use the median as it has a higher breakdown point and is therefore more accurate than regular Z -scores (Rousseeuw and Hubert 2011). The cause for each identified outlier was carefully investigated before any value could be excluded from the dataset (Table S1).

All data were analyzed with MetaWin v. 2.1.5 (Sinauer Associates Inc. 2000), and graphs for visual illustrations were created using the graphic data tool DataGraph v. 3.1.2 (VisualDataTools Inc. 2013).

3 Results

3.1 Publication bias

For each of the three time periods considered in this study (prior to IPCC-AR4 2007, after IPCC-AR4 2007 and before Climategate 2009, and after Climategate 2009), the funnel plots showed no evidence of statistically non-significant results being under-represented (Fig. 1); there were no holes around $Y = 0$, nor were there conspicuous gaps or holes in other parts of the funnels (Fig. 1). Strong fail-safe sample sizes confirmed that the effect sizes were robust and that publication bias was not detected (Rosenthal 1979). We further tested the robustness of results by re-sampling single results from articles and reproducing funnel plots (see Supplementary Material Fig. S4 a–j).

3.2 Number of studies, effect size, and abstract versus main

The number of articles about climate change in ocean ecosystems has increased annually since 1997, peaking within 2 years after IPCC 2007 and subsiding after Climategate 2009 (Fig. 2).

Before Climategate, reported effect sizes were significantly larger in article abstracts than in the main body of articles, suggesting a systematic bias in how authors are communicating results in scientific articles: Large, significant effects were emphasized where readers are most likely to see them (in abstracts), whereas small or non-significant effects were more often found in the technical results sections where we presume they are less likely to be seen by the majority of readers, especially non-scientists. Moreover, between IPCC 2007 and Climategate, when publication rates on ocean climate change were greatest, the difference in effect sizes reported in the abstract and body of reports was also greatest (Fig. 2). After Climategate, publication rates about ocean climate change fell, the magnitude of reported effect sizes in abstracts diminished, and the difference in effect sizes between abstracts and the body of reports returned to a level comparable to pre-IPCC 2007 (Fig. 2).

3.3 Impact factor

Journals with an impact factor greater than 9 published significantly larger effect sizes than journals with an impact factor of less than 9 (Fig. 3). Regardless of the impact factor, journals reported significantly larger effect sizes in abstracts than in the main body of articles; however, the difference between mean effects in abstracts versus body of articles was greater for journals with higher impact factors. We also detected a small, yet statistically significant, negative relationship between reported sample size and journal impact factor, which was largely driven by the large effects reported in high impact factor journals (Fig. 4). Despite the larger effect sizes, journals with high impact factors published results with generally lower sample sizes.

4 Discussion

Our meta-analysis did not find evidence of small, statistically non-significant results being under-reported in our sample of climate change articles. This result opposes findings by Michaels (2008) and Reckova and Irsova (2015), which both found publication bias in the

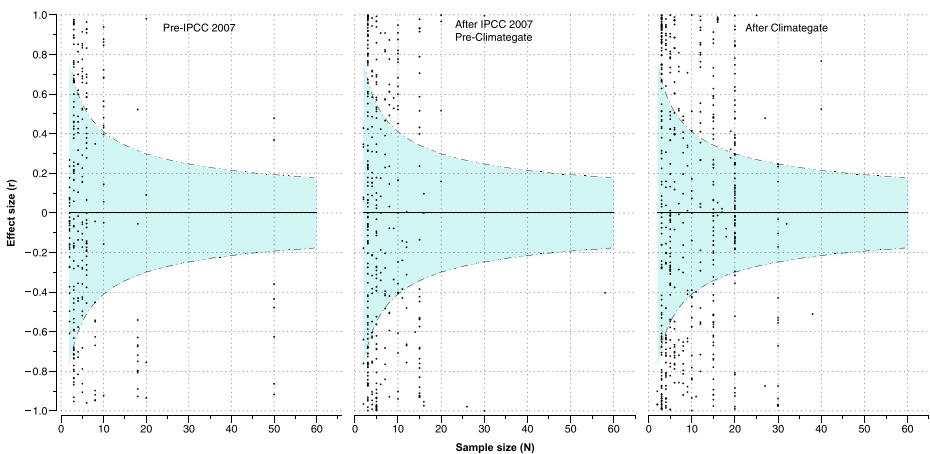


Fig. 1 Funnel plots representing effect size (r) as a function of sample size (N). The shaded areas represent results that were not significant statistically for the main dataset. **a** Pre-IPCC 2007 ($n = 265$). **b** After IPCC 2007/pre-Climategate ($n = 345$). **c** After Climategate ($n = 544$). n denotes number of experiments

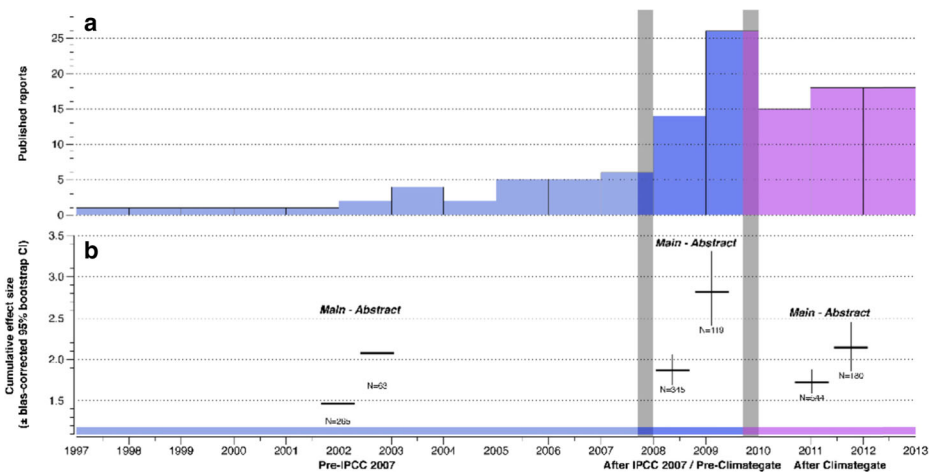


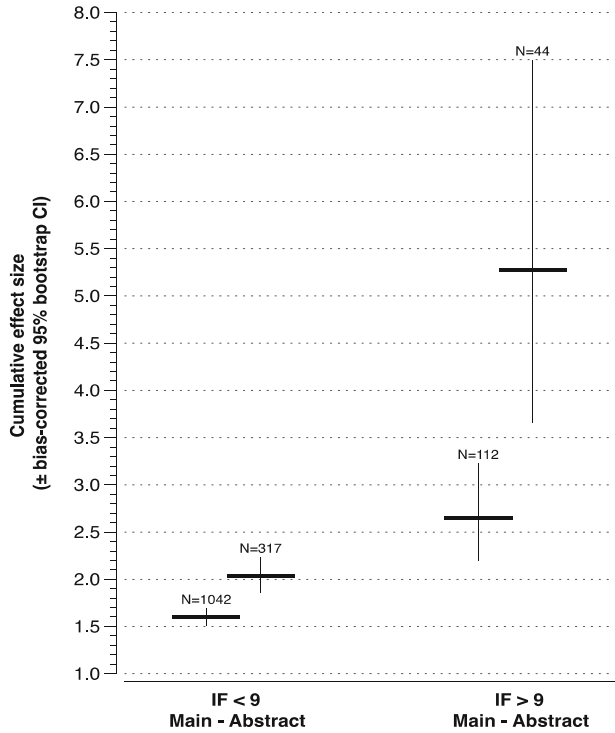
Fig. 2 Publication rate. **a** Number of published reports for each year. The *two vertical grey bars* illustrate the timing of IPCC 2007 and Climategate 2009. **b** Cumulative effect sizes of Hedges' d and bias-corrected 95% bootstrap confidence intervals for the magnitude of climate-change effects. Mean effect sizes are computed separately for results presented in abstracts and in the main body of articles. N sample size. Pre-IPCC 2007 main dataset: $d = 1.46$; $CI = 1.30\text{--}1.63$; $df = 264$, $FSN = 36,299$. Abstract dataset: $d = 2.08$; $CI = 1.73\text{--}2.51$; $df = 62$, $FSN = 3,475$; $P < 0.05$. After IPCC 2007/pre-Climategate main dataset: $d = 1.87$; $CI = 1.69\text{--}2.06$; $df = 344$, $FSN = 79,576$. Abstract dataset: $d = 2.82$; $CI = 2.41\text{--}3.31$; $df = 118$, $FSN = 11,557$; $P < 0.05$. After Climategate main dataset: $d = 1.72$; $CI = 1.59\text{--}1.88$; $df = 543$, $FSN = 214,674$. Abstract dataset: $d = 2.14$; $CI = 1.85\text{--}2.46$; $df = 176$, $FSN = 26,480$; $P = n.s.$ d Hedges' d , CI bias-corrected 95% confidence intervals, df degrees of freedom (one less than total sample), FSN fail-safe numbers. P , $n.s.$ probability that abstract results and main text results differ

global climate change literature, albeit with a smaller sample size for their meta-analysis and in other sub-disciplines of climate change science. Michaels (2008) examined articles from Nature and Science exclusively, and therefore, his results were influenced strongly by the editorial position of these high impact factor journals with respect to reporting climate change issues. We believe that the results presented here have added value because we sampled a broader range of journals, including some with relatively low impact factor, which is probably a better representation of potential biases across the entire field of study. Moreover, several end users and stakeholders of science, including other scientists and public officials, base their research and opinions on a much broader suite of journals than Nature and Science.

However, our meta-analysis did find multiple lines of evidence of biases within our sample of articles, which were perpetuated in journals of all impact factors and related largely to how science is communicated: The large, statistically significant effects were typically showcased in abstracts and summary paragraphs, whereas the lesser effects, especially those that were not statistically significant, were often buried in the main body of reports. Although the tendency to isolate large, significant results in abstracts has been noted elsewhere (Fanelli 2012), here we provide the first empirical evidence of such a trend across a large sample of literature.

We also discovered a temporal pattern to reporting biases, which appeared to be related to seminal events in the climate change community and may reflect a socio-economic driver in the publication record. First, there was a conspicuous rise in the number of climate change publications in the 2 years following IPCC 2007, which likely reflects the rise in popularity (among public and funding agencies) for this field of research and the increased appetite among journal editors to publish these articles. Concurrent with increased publication rates

Fig. 3 Cumulative effect sizes of Hedges' *d* and bias-corrected 95% bootstrap confidence intervals for the magnitude of climate change effects for journals with impact factor above or below 9. Results are computed separately for data from abstracts and the main body of reports. *N* denotes the sample size. IF < 9 main dataset: *d* = 1.60; CI = 1.51–1.69; df = 1042, FSN = 696,107, *P* < 0.05. Abstract dataset: *d* = 2.04; CI = 1.86–2.24; df = 316, FSN = 83,671, *P* < 0.05. IF > 9 main dataset: *d* = 2.65; CI = 2.20–3.23; df = 111, FSN = 10,131, *P* < 0.05. Abstract dataset: *d* = 5.27; CI = 3.66–7.50; df = 44, FSN = 2298, *P* < 0.05. Abbreviations as in Fig. 2 legend



was an increase in reported effect sizes in abstracts. Perhaps a coincidence, the apparent popularity of climate change articles (i.e., number of published articles and reported effect sizes) plummeted shortly after Climategate, when the world media focused its scrutiny on this field of research, and perhaps, popularity in this field waned (Fig. 1). After Climategate,

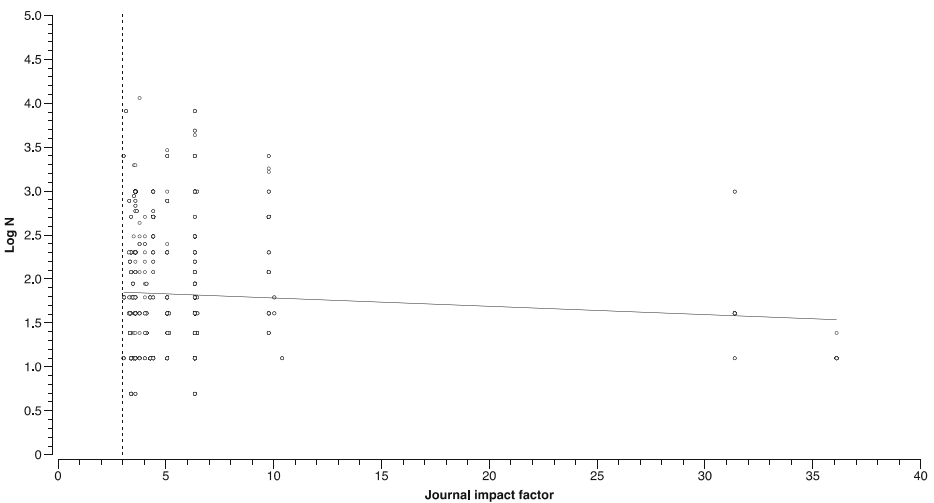


Fig. 4 Relationship between journal impact factor and sample size for experimental results ($R^2 = 0.004$; $P < 0.05$)

reported effect sizes also dropped, as did the difference in effects reported in abstracts versus main body of articles. The positive effect we see post IPCC 2007, and the negative effect post Climategate, may illustrate a combined effect of editors' or referees' publication choices and researchers' propensity to submit articles or not. However, since meta-analysis is correlative, it does not elucidate the mechanisms underlying observed patterns.

Similar stylistic biases were found when comparing articles from journals with high impact factors to those with low impact factors. High impact factors were associated with significantly larger reported effect sizes (and lower sample sizes; see Fig. 4); these articles also had a significantly larger difference between effects reported in abstracts versus the main body of their reports (Fig. 3). This trend appears to be driven by a small number of journals with large impact factors; however, the result is consistent with those of supplementary studies. For example, our results corroborate with others by showing that high impact journals typically report large effects based on small sample sizes (Fraley and Vazire 2014), and high impact journals have shown publication bias in climate change research (Michaels 2008, and further discussed in Radetzki 2010).

Stylistic biases are less concerning than a systematic tendency to under-report non-significant effects, assuming researchers read entire reports before formulating theories. However, most audiences, especially non-scientific ones, are more likely to read article abstracts or summary paragraphs only, without perusing technical results. The onus to effectively communicate science does not fall entirely on the reader; rather, it is the responsibility of scientists and editors to remain vigilant, to understand how biases may pervade their work, and to be proactive about communicating science to non-technical audiences in transparent and un-biased ways. Ironically, articles in high impact journals are those most cited by other scientists; therefore, the practice of sensationalizing abstracts may bias scientific consensus too, assuming many scientists may also rely too heavily on abstracts during literature reviews and do not spend sufficient time delving into the lesser effects reported elsewhere in articles.

Despite our sincerest aim of using science as an objective and unbiased tool to record natural history, we are reminded that science is a human construct, often driven by human needs to tell a compelling story, to reinforce the positive, and to compete for limited resources—publication trends and communication bias is a proof of that.

Acknowledgements We are grateful to Roger Butlin, Christer Brönmark, A. Richard Palmer, Tobias Uller, Charlie Cornwallis and Johannes Persson for constructive advice, and a special thanks to Dean Adams. We are also thankful to Michael MacAskill, Jessica Gurevitch and Shinichi Nakagawa for technical advice. JH was funded by a Marie Curie European Reintegration Grant (PERG08-GA-2010-276915) and by the Royal Physiographic Society in Lund.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Adams D (2010) “Climategate” review clears scientists of dishonesty over data. *The Guardian*, Wednesday 7 July

- Adams DC, Gurevitch J, Rosenberg MS (1997) Resampling tests for meta-analysis of ecological data. *Ecology* 78:1277–1283. doi:10.1890/0012-9658(1997)078[1277:RTFMAO]2.0.CO;2
- Anderegg WRL, Prall JW, Harold J, Schneider SH (2010) Expert credibility in climate change. *Proc Natl Acad Sci U S A* 107:12107–12109. doi:10.1073/pnas.1003187107
- Begg CB, Mazumdar M (1994) Operating characteristics of a rank correlation test for publication bias. *Biometrics* 50:1088–1101. doi:10.2307/2533446
- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd
- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2010) A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Meth* 1:97–111. doi:10.1002/jrsm.12
- Carlsson-Kanyama A, Hörnsten Friberg L (2012) Views on climate change and adaptation among politicians and directors in Swedish municipalities. FOI-R–3441–SE. FOI Totalförsvarets forskningsinstitut, FOI, Stockholm
- Cooper H (1998) *Synthesizing research: a guide for literature reviews* (3rd ed.) Sage Thousand Oaks
- Corner AJ, Whitmarsh LE, Xenias D (2012) Uncertainty, scepticism and attitudes towards climate change: biased assimilation and attitude polarisation. *Clim Chang* 114:463–478. doi:10.1007/s10584-012-0424-6
- DataGraph v. 3.1.2. (2013) Adalsteinsson D, VisualDataTools Inc
- Doran PT, Kendall Zimmerman M (2009) Examining the scientific consensus on climate change. *Science* 90:22–23. doi:10.1029/2009EO030002
- Fanelli D (2012) Negative results are disappearing from most disciplines and countries. *Scientometrics* 90:891–904. doi:10.1007/s11192-011-0494-7
- Fanelli D (2013) Positive results receive more citations, but only in some disciplines. *Scientometrics* 94:701–709. doi:10.1007/s11192-012-0757-y
- Fraley RC, Vazire S (2014) The N-pact factor: evaluating the quality of empirical journals with respect to sample size and statistical power. *PLoS ONE* 9:e109019. doi:10.1371/journal.pone.0109019
- GraphClick v. 3.0. Retrieved 2011-08-23. Arizona Software. See: www.arizona-software.ch
- Gurevitch J, Hedges LV (1993) Meta-analysis: combining the results of independent experiments. In: Scheiner SM, Gurevitch J (eds) *Design and analysis of ecological experiments*. Chapman & Hall, New York, p 445
- Hedges LV, Olkin I (1985) *Statistical Methods for Meta-Analysis*. Academic Press, New York
- Holland D (2007) Bias and concealment in the IPCC process: the “hockey-stick” affair and its implications. *Energy Environ* 18:951–983. doi:10.1260/095830507782616788
- Hollander J (2008) Testing the grain-size model for the evolution of phenotypic plasticity. *Evolution* 62:1381–1389. doi:10.1111/j.1558-5646.2008.00365.x
- Idso C, Singer FS (2009) Climate change reconsidered: 2009 report of the nongovernmental panel on climate change (NIPCC). The Heartland Institute, Chicago
- IPCC (2007) In: Core Writing Team, Pachauri RK, Reisinger A (eds) *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, p 104
- Kicinski M (2013) Publication bias in recent meta-analyses. *PLoS ONE* 8:e81823. doi:10.1371/journal.pone.0081823
- Kicinski M, Springate DA, Kontopantelis E (2015) Publication bias in meta-analyses from the Cochrane database of systematic reviews. *Stat Med* 34:2781. doi:10.1002/sim.6525
- Light RJ, Pillemer DB (1984) *Summing up: The science of reviewing research*. Cambridge: Harvard Univ. Press
- Michaels PJ (2008) Evidence for “Publication bias” concerning global warming in Science and Nature. *Energy Environ* 19:287–301. doi:10.1260/095830508783900735
- Michaels PJ (2010) The Climategate whitewash continues. *Wall Street J*
- Palmer RA (1999) Notes and comments—detecting publication bias in meta-analyses: a case study of fluctuating asymmetry and sexual selection. *Am Nat* 154:220–233. doi:10.1086/303223
- Palmer RA (2000) Quasireplication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annu Rev Ecol Syst* 31:441–480. doi:10.1146/annurev.ecolsys.31.1.441
- Radetzki M (2010) The fallacies of concurrent climate policy efforts. *Ambio*. 2010 May; 39(3): 211–222. Published online 2010 Jun 3. doi: 10.1007/s13280-010-0029-0
- Reckova D, Irsova Z (2015) Publication bias in measuring anthropogenic climate change. *Energy Environ* 26: 853–862. doi:10.1260/0958-305X.26.5.853
- Rosenberg MS (2005) The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* 59:464–468. doi:10.1111/j.0014-3820.2005.tb01004.x
- Rosenberg MS, Adams DC, Gurevitch J (2000) *MetaWin*. Statistical software for meta-analysis. Sinauer Associates, Inc., v. 2.1.5. Sunderland, Massachusetts; See: <http://www.metawinsoft.com/>.
- Rosenthal R (1979) The “file drawer problem” and tolerance for null results. *Psych Bull* 86:638–641. doi:10.1037/0033-2909.86.3.638

- Rousseeuw PJ, Hubert M (2011) Robust statistics for outlier detection. *WIREs Data Min Knowl Discovery* 1:73–79. doi:[10.1002/widm.2](https://doi.org/10.1002/widm.2)
- Shadish WR, Haddock CK (1994) Combining estimates of effect size. In: Cooper H, Hedges LV (eds) *The Handbook of research synthesis*. Russell Sage, New York
- The Copenhagen Diagnosis (2009) In: Allison I et al (eds) *Updating the world on the latest climate science*. The University of New South Wales Climate Change Research Centre (CCRC), Sydney, p 60