

Comment on 'Are physicists afraid of mathematics?'

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 New J. Phys. 18 118003

(<http://iopscience.iop.org/1367-2630/18/11/118003>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 137.222.249.190

This content was downloaded on 21/11/2016 at 10:52

Please note that [terms and conditions apply](#).

You may also be interested in:

[Are physicists afraid of mathematics?](#)

Jonathan E Kollmer, Thorsten Pöschel and Jason A C Gallas

[Reply to Comment on 'Are physicists afraid of mathematics?'](#)

Jonathan E Kollmer, Thorsten Pöschel and Jason A C Gallas

[arXiv:astro-ph Position Dependent Citation Rates](#)

J. P. Dietrich

[Letters that made a journal](#)

Michael Banks

[Nuclear Physics: A fifth force of nature?](#)

W Gough

[Particle Physics: Neutrino oscillations: you decide](#)

Matin Durrani

[The first-mover advantage in scientific publication](#)

M. E. J. Newman



COMMENT

Comment on ‘Are physicists afraid of mathematics?’

OPEN ACCESS

RECEIVED

16 February 2016

REVISED

4 April 2016

ACCEPTED FOR PUBLICATION

21 October 2016

PUBLISHED

11 November 2016

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Andrew D Higginson¹ and Tim W Fawcett

Centre for Research in Animal Behaviour, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QG, UK

¹ Author to whom any correspondence should be addressed.E-mail: a.higginson@exeter.ac.uk and T.W.Fawcett@exeter.ac.uk**Abstract**

In 2012, we showed that the citation count for articles in ecology and evolutionary biology declines with increasing density of equations. Kollmer *et al* (2015 *New J. Phys.* **17** 013036) claim this effect is an artefact of the manner in which we plotted the data. They also present citation data from *Physical Review Letters* and argue, based on graphs, that citation counts are unrelated to equation density. Here we show that both claims are misguided. We identified the effects in biology not by visual means, but using the most appropriate statistical analysis. Since Kollmer *et al* did not carry out any statistical analysis, they cannot draw reliable inferences about the citation patterns in physics. We show that when statistically analysed their data actually do provide evidence that in physics, as in biology, citation counts are lower for articles with a high density of equations. This indicates that a negative relationship between equation density and citations may extend across the breadth of the sciences, even those in which researchers are well accustomed to mathematical descriptions of natural phenomena. We restate our assessment that this is a genuine problem and discuss what we think should be done about it.

1. Mathematics plays a vital role in the sciences

Mathematical theory is an indispensable part of scientific research, capturing the essence of fundamental physical, chemical and biological processes with greater clarity, precision, rigor, and brevity than verbal arguments can achieve. In a range of disciplines, efficient dialogue between theoretical developments and empirical testing is critical to driving science forwards [1–8]. Reports of a barrier to communication between theoretical and empirical research [9–11] should therefore arouse concern.

In a recent study [12], we showed that the citation counts of articles in ecology and evolutionary biology are negatively associated with the density of mathematical equations presented in the main text. In a paper published in the *New Journal of Physics*, Kollmer *et al* [13]—hereafter KPG—attacked our interpretation of the citation patterns in ecology and evolutionary biology, criticized our methodology and argued that there is no evidence for a negative impact of high equation density on citation counts. Responding to our call for similar analyses in other fields [14], KPG also investigated the relationship between equation density and citation count in a leading multidisciplinary physics journal, *Physical Review Letters*. As they did for the biology data, they argued that there was no evidence for a relationship. Here we show that the conclusions drawn by KPG are incorrect.

2. Our original analysis is objective and valid

KPG present our data in alternative graphical formats (their figure 1) and suggest that the effect we found is simply an artefact of how we binned the data. This is incorrect: our conclusions were based on a formal statistical analysis that did not involve any binning of data, but instead treated equation density as a continuous covariate. This statistical analysis—a generalized linear model (GLM) with a negative binomial error function—is appropriate for count data that are extremely over-dispersed (in this case, the most cited papers are much more

highly cited than expected from a Poisson distribution). We used binning merely to illustrate the patterns graphically, as a visual aid to readers' intuition. This binning was entirely separate from our statistical analysis, so the claim that this 'strongly influences the final outcome of the analysis' (p 3) is unfounded.

KPG argue that 'the citation data is so noisy that its [sic.] not reliable for identifying unambiguous trends capable of predicting citation success' (p 7). Yet scientists in many fields routinely analyse noisy data using statistical modelling. In this case, the noisiness of the patterns implies that citation rates are, reassuringly, more strongly influenced by a combination of factors other than equation density, which alone explains a relatively small proportion of the variance—a point we emphasised in our original paper [12]. Despite incorporating all of this noisiness, our statistical models consistently revealed a significant (and sizeable) negative effect of equation density on citation count.

KPG claim that our 'original presentation of the data ... disregarded completely the fact that by far the two most cited papers ... have very high equation densities and high citation counts coming from non-theoretical papers' (p 3). This is incorrect: both our graphical representations and the formal statistical analysis included all data points, regardless of whether they were 'outliers'. Statistically, nothing can be reliably inferred from two examples of heavily cited, equation-dense outliers selected from a sample of 649 articles. These two outliers have exceptional citation counts because of factors other than equation density (e.g. exceptional scientific importance); counts that may have been even higher if the papers were less equation-dense.

Finally, KPG conclude that our findings are 'strongly dependent on their artificial subdivision of papers into theoretical and non-theoretical work' (p 6). This is not true because the effect is still present when analysing the entire sample (22% drop in citations for each additional equation per page), without any subdivision into theoretical and non-theoretical papers.

3. Citations in physics show a negative effect of equation density

KPG presented data on the number of equations and citations for the set of papers published in volumes 94 and 104 of *Physical Review Letters*. They reported no relationship between equation density and number of citations but they did not carry out a formal statistical analysis, so their conclusions are debatable: visual comparison of binned data is a subjective and unreliable way to infer statistical trends in continuous data. In two of their figures (figures 2(a) and (d)) KPG noted a slight decrease in citations for more equation-dense papers, but stated that this 'decrease is not significant because it is well within the large error bars' (p 5). However, drawing statistical conclusions based on error bars is fraught with difficulty [15], particularly when the error bars are descriptive (e.g. standard deviation, as used by KPG) rather than inferential (e.g. standard error of the mean, SEM, as used by [12]) [16]. The degree of overlap between error bars may be a useful guide when exploring data to assess whether group means are more different than would be expected by chance [16], provided that one divides the data into only a few groups. As continuous data is binned into more and more groups, error bars will tend to become larger simply because the sample each is based on gets smaller. The size of error bars for small bins will be 'sometimes dominated by a single high-impact paper' (KPG, p 5), and thus the error bars become less informative of overall trends.

To assess the effect of equation density objectively we carried out the same type of statistical analysis as in our original paper, using the data set presented by KPG (for details see doi: [10.5281/zenodo.58792](https://doi.org/10.5281/zenodo.58792)). This formal and robust analysis shows that equation density has a statistically significant negative effect on the number of citations, leading on average to 6% fewer citations for each additional equation per page. This effect increases to 8% for papers that have been cited fewer than 100 times each.

4. Our viewpoint has been misinterpreted

KPG state that they were motivated to write their article by our 'surprising attempt of blaming math ... for a lack of success in getting citations' (p 2). This was certainly not our intention: both of us regularly publish papers describing mathematical and computational models, so we sincerely hope that a low citation rate for mathematical work is not inevitable. Rather, we suggested that an immediate, pragmatic solution to this apparent problem would be to reduce the density of equations and add explanatory text for non-specialised readers. We have repeatedly emphasized in our papers on this topic [12, 14] that it is equation *density*, not the *number* of equations, that is associated with citation counts. Thus, we recommend that the authors of theoretical work use more text to explain their equations clearly.

Our suggestions for how the negative impact of equation density can be remedied are unchanged: essential equations capturing the assumptions and structure of a model should be presented in the main text, whereas non-essential equations, such as those describing intermediate steps to solutions, need only be given in the appendices. This would have the effect of reducing equation density, and allow more space to explain the

assumptions and implications of the work underlying the essential equations. Making such adjustments to presentation could potentially have strong beneficial effects on citation patterns. For instance, our statistical model predicts that, all else being equal, the 45 articles in KPG's data that are moderately well cited (50–100 citations) and equation dense (≥ 2 equations per page) would have attracted an additional 476 citations (17% of their total) if the authors had halved the density of equations in the main text.

The finding that the negative effect of equation density is also present in a data set from physics suggests that the phenomenon is not restricted to the life sciences, but extends to fields with a traditionally greater reliance on mathematics. The effect we found was considerably weaker than in a sample of papers from ecology and evolutionary biology [12], but still sizeable and statistically significant. This suggests that the problem may be even more widespread than we originally thought, perhaps affecting all disciplines that rely on mathematics to understand natural phenomena.

Ideally, the impact of scientific work should be determined by its scientific merit, rather than by presentational style. Unfortunately, it is clear that scientifically strong papers may have reduced impact if not presented in an accessible manner. We reiterate our view that all scientists aiming to communicate theory in the most effective way should take this issue seriously, rather than claiming it does not exist.

Acknowledgments

This work was supported by a European Research Council Advanced Grant (250209) awarded to Alasdair Houston and NERC Independent Research Fellowship (NE/L011921/1) awarded to ADH. We are indebted to the American Physical Society and *Physical Review Letters* for granting access to their article data and Jonathan Kollmer, Thorsten Pöschel and Jason Gallas for supplying it.

References

- [1] Lunney D, Pearson J M and Thibault C 2003 *Rev. Mod. Phys.* **75** 1021
- [2] Marcus R A 1993 *Rev. Mod. Phys.* **65** 599
- [3] Deco G, Jirsa V K, Robinson P A, Breakspear M and Friston K 2008 *PLoS Comput. Biol.* **4** e1000092
- [4] Elena S F and Froissart R 2010 *Philos. Trans. R. Soc. Lond. B* **365** 1867
- [5] Kareiva P 1989 *Perspectives in Ecological Theory* ed J Roughgarden, R M May and S A Levin (Princeton, NJ: Princeton University Press) pp 68–88
- [6] Raupach M R *et al* 2005 *Glob. Change. Biol.* **11** 378
- [7] Cohen J E 2004 *PLoS Biology* **2** e439
- [8] Servideo M R, Brandvain Y, Dhole S, Fitzpatrick C L, Goldberg E E, Stern C A, van Cleve J and Yeh D J 2014 *PLoS Biology* **12** e1002017
- [9] Butlin R K and Tregenza T 2005 *J. Evol. Biol.* **18** 1205
- [10] Scheiner S M 2005 *Ecol. Lett.* **16** 1421
- [11] Haller B C 2014 *BioScience* **64** 907
- [12] Fawcett T W and Higginson A D 2012 *Proc. Natl Acad. Sci. USA* **109** 11735
- [13] Kollmer J E, Pöschel T and Gallas J A C 2015 *New J. Physics* **17** 013036
- [14] Fawcett T W and Higginson A D 2012 *Proc. Natl Acad. Sci. USA* **109** E3058
- [15] Belia S, Fidler F, Williams J and Cumming G 2005 *Psychological Methods* **10** 389
- [16] Cumming G, Fidler F and Vaux D L 2007 *J. Cell Biol.* **177** 7