

# Heavy use of equations impedes communication among biologists

Tim W. Fawcett<sup>1</sup> and Andrew D. Higginson

School of Biological Sciences, University of Bristol, Bristol BS8 1UG, United Kingdom

Edited<sup>†</sup> by Robert M. May, University of Oxford, Oxford, United Kingdom, and approved June 6, 2012 (received for review April 4, 2012)

**Most research in biology is empirical, yet empirical studies rely fundamentally on theoretical work for generating testable predictions and interpreting observations. Despite this interdependence, many empirical studies build largely on other empirical studies with little direct reference to relevant theory, suggesting a failure of communication that may hinder scientific progress. To investigate the extent of this problem, we analyzed how the use of mathematical equations affects the scientific impact of studies in ecology and evolution. The density of equations in an article has a significant negative impact on citation rates, with papers receiving 28% fewer citations overall for each additional equation per page in the main text. Long, equation-dense papers tend to be more frequently cited by other theoretical papers, but this increase is outweighed by a sharp drop in citations from nontheoretical papers (35% fewer citations for each additional equation per page in the main text). In contrast, equations presented in an accompanying appendix do not lessen a paper's impact. Our analysis suggests possible strategies for enhancing the presentation of mathematical models to facilitate progress in disciplines that rely on the tight integration of theoretical and empirical work.**

impact factor | mathematical formula | mathematical literacy | theoretical biology

The efficient exchange of new findings and insights between empirical and theoretical approaches is critical to a range of scientific disciplines, including nuclear physics (1), physical chemistry (2), neuroscience (3), epidemiology (4), ecology (5), and atmospheric science (6). In evolutionary biology, for example, the integration of empirical and theoretical work is essential for understanding how natural selection shapes organisms and their interactions (7–16). Most biological research is empirical, yet empirical studies rely fundamentally on theory for generating testable predictions and interpreting observations. In return, empirical data provide both tests of established theory and guidance in the development of new models.

However, the importance of presenting theory in sufficient technical detail can sometimes conflict with the need to communicate the essence of a model in a clear, accessible manner. Concise and precise description of the structure of a mathematical model demands the use of equations, but such technical details might deter a broad audience of scientists doing largely empirical research. A cursory reading of the biological literature reveals that many empirical studies build largely on other empirical studies, with little direct reference to relevant theory. This observation suggests a breakdown of communication that may impede scientific progress.

To explore the extent of this problem, we systematically investigated how the use of mathematical equations affects the scientific impact of studies in ecology and evolution. We examined the use of equations and obtained citation data for all papers (total  $n = 649$ ; Dataset S1) published in 1998 in the top three journals specializing in ecology and evolution: *Evolution*, *Proceedings of the Royal Society of London B*, and *The American Naturalist*. We find that heavy use of equations reduces citation rates, because papers with a high density of equations per page attract fewer citations from nontheoretical papers. Our results suggest possible strategies

for enhancing the presentation of mathematical models to facilitate progress in disciplines that rely on the tight integration of theoretical and empirical work.

## Results

To quantify the technical level of any theory presented in the articles, we counted equations, inequalities, and other mathematical expressions (hereafter referred to simply as “equations”) in the main text and any printed appendixes. We divided this count by the number of pages to give a measure of equation density, which ranged from 0 to 7.29 equations per page (mean  $\pm$  SEM:  $0.43 \pm 0.04$ ) and was uncorrelated with the length of the article ( $r_{647} = 0.056$ ,  $P = 0.151$ ). To assess impact, we obtained citation data for these articles from the Science Citation Index Expanded on the Thomson Reuters Web of Science in May 2011, excluding any self-citations (i.e., citing papers for which one or more of the author surnames matched one or more of the author surnames for the cited paper). The number of citations varied widely, ranging from 0 to 374 with a mean  $\pm$  SEM of  $44.80 \pm 1.98$  citations (excluding self-citations). Controlling for a significant positive effect of paper length (Table 1, *All citations*), the use of equations has a striking influence on this measure of impact. Equation density negatively affects citation rates, leading on average to 22% fewer citations for each additional equation per page (Table 1, *All citations*).

We might expect this effect to be driven largely by a reduction in nontheoretical citations. To investigate this hypothesis, we searched for the term “model\*” (excluding some common empirical uses such as “experimental model\*”) in the title or abstract of the citing articles and used the presence of this term as a proxy for whether the citing paper was a theoretical one. This search identified 6,229 (22.2%) of the 28,068 citing articles as “theoretical.” We validated our proxy by examining a randomly selected subset of 200 citing articles, which showed that 84.5% were correctly classified as theoretical or nontheoretical. As expected, the negative effect of equation density is strongest for nontheoretical papers, which provide 27% fewer citations for each additional equation per page (Table 1, *Nontheoretical citations*). Articles less than 10 pages long with up to 0.5 equations per page are just as well cited as those with no equations, but increasing the equation density to more than one equation per page more than halves the number of nontheoretical citations (Fig. 1A). In contrast, longer papers (>9 pages) receive more citations when they are completely equation-free, but beyond this difference, there appears to be no effect of quantitative changes in equation density (Fig. 1A). Statistically, however, the effect of equation density on nontheoretical citations was consistent across papers of different lengths (nonsignificant interaction term; Table 1, *Nontheoretical citations*).

Author contributions: T.W.F. and A.D.H. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

<sup>†</sup>This Direct Submission article had a prearranged editor.

<sup>†</sup>To whom correspondence should be addressed. E-mail: tim.fawcett@cantab.net.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1205259109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1205259109/-DCSupplemental).

**Table 1. Variables affecting the number of citations by all papers, nontheoretical papers, and theoretical papers**

Parameter	All citations			Nontheoretical citations			Theoretical citations		
	OR (95% CI)	Wald z	P	OR (95% CI)	Wald z	P	OR (95% CI)	Wald z	P
Intercept	28.67 (20.69–39.74)	20.189	<0.001	20.93 (14.77–29.66)	17.135	<0.001	6.14 (4.17–9.03)	9.219	<0.001
Density of equations per page	<b>0.78 (0.66–0.93)</b>	<b>−2.782</b>	<b>0.005</b>	<b>0.73 (0.61–0.88)</b>	<b>−3.244</b>	<b>0.001</b>	0.97 (0.79–1.18)	−0.338	0.735
Total no. of pages	<b>1.05 (1.02–1.07)</b>	<b>3.929</b>	<b>&lt;0.001</b>	<b>1.05 (1.02–1.08)</b>	<b>3.692</b>	<b>&lt;0.001</b>	<b>1.05 (1.02–1.08)</b>	<b>3.379</b>	<b>0.001</b>
Published in <i>Evolution</i> (cf. <i>Am. Nat.</i> )	0.95 (0.76–1.18)	−0.494	0.622	1.07 (0.85–1.35)	0.573	0.567	<b>0.70 (0.54–0.91)</b>	<b>−2.692</b>	<b>0.007</b>
Published in <i>Proceedings B</i> (cf. <i>Am. Nat.</i> )	1.14 (0.90–1.43)	1.102	0.270	1.22 (0.95–1.55)	1.565	0.118	0.93 (0.71–1.21)	−0.561	0.575
Equation density × no. of pages	1.02 (1.00–1.04)	1.636	0.102	1.01 (0.99–1.03)	0.937	0.349	<b>1.03 (1.01–1.05)</b>	<b>2.443</b>	<b>0.015</b>

The table shows statistical results from a generalized linear model with a negative binomial error structure. For a unit increase in the explanatory variable, the number of citations changes by a factor given by the OR, shown here with a 95% CI. For example, an OR of 0.78 implies a decrease of 22%, whereas an OR of 1.05 implies an increase of 5%. Significant effects ( $P < 0.05$ ) based on the Wald z statistic are highlighted in bold.

Controlling for a significant effect of the journal of publication, there was no main effect of equation density on citations by theoretical papers (Table 1, *Theoretical citations*). We did, however, record a significant positive interaction between equation density and the length of the cited paper. This interaction occurs because papers of 10 pages or more have increased citation success when they contain more than 0.5 equations per page (Fig. 1*B*), implying that long, equation-dense papers are more likely to be cited by other papers presenting theoretical work.

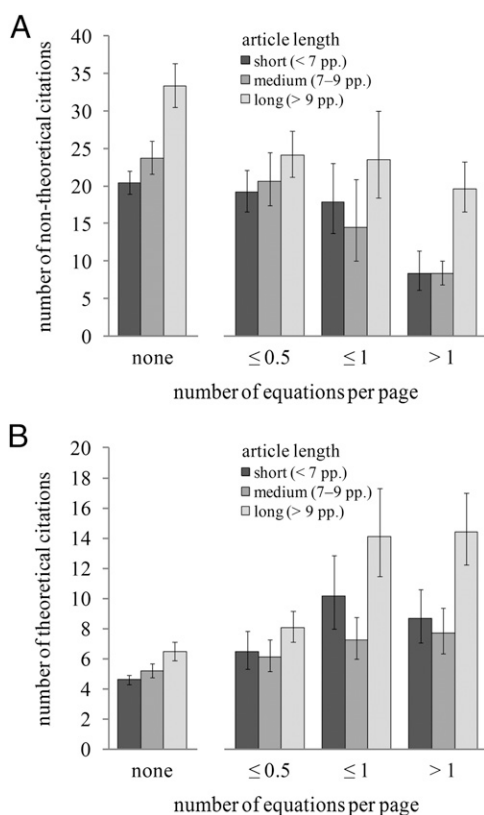
Next, we distinguished between equations presented in the main text and those presented in an appendix. The overall number of citations decreases with the density of equations in the main text,

each additional equation per page leading to a 28% drop in citations (Table 2, *All citations*). In contrast, equations presented in an appendix have no impact on citation rates (Table 2, *All citations*). Again these effects are largely driven by citation patterns in the nontheoretical literature. Citations by nontheoretical papers decrease by 35% for each additional equation per page presented in the main text (Table 2, *Nontheoretical citations*). For papers less than 10 pages long, the citation count more than halves when the main-text equation density is increased from 0.5 or less to more than one per page (Fig. 2*A*), whereas for longer papers (>9 pages), any equations in the main text appear to reduce citation success. Additional equations in the appendix, however, have no effect on nontheoretical citation rates (Table 2, *Nontheoretical citations* and Fig. 2*B*). Citations by theoretical papers are unaffected by the density of equations in either the main text or the appendixes (Table 2, *Theoretical citations*), but the interaction between the density of main-text equations and the length of the paper was close to significance ( $P = 0.074$ ), again suggesting that long, equation-dense articles garner more citations from other theoretical papers.

The above findings suggest that these effects are not merely due to papers containing some equations being generally less well cited than those containing none. To check whether this interpretation is correct, we restricted our sample of cited papers to those containing at least one equation ( $n = 247$ ). This analysis yielded similar results: The overall number of citations goes down with increasing equation density [odds ratio (OR) = 0.78, 95% confidence interval (CI) = 0.64–0.96, Wald  $z = -2.393$ ,  $P = 0.017$ ], and this effect is due to equations in the main text (OR = 0.72, 95% CI = 0.55–0.93, Wald  $z = -2.514$ ,  $P = 0.012$ ) rather than equations in the appendixes (OR = 1.01, 95% CI = 0.67–1.52, Wald  $z = 0.042$ ,  $P = 0.966$ ). Thus, there is a quantitative effect of increasing the density of equations, not simply an aversion to citing papers containing any mathematics.

## Discussion

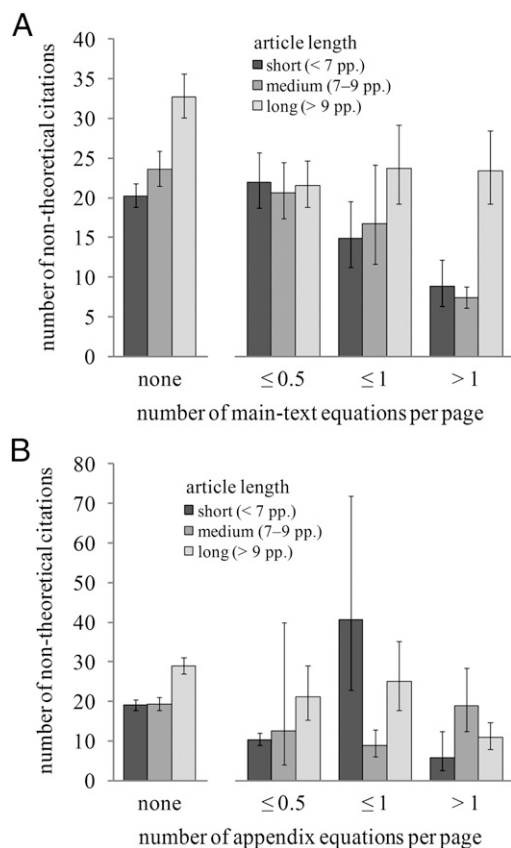
A paper's impact ought to be determined largely by its scientific merit, in terms of its novelty, rigor, breadth of interest, and other aspects of quality that are difficult or impossible to assess objectively, rather than by the particular way in which the methodology is presented. However, our results suggest that a scientifically strong theoretical paper risks dramatically reducing its impact by presenting its mathematical details in a highly technical manner. Long and equation-dense papers tend to be better cited by others doing theoretical work—perhaps because such papers offer the most in-depth theoretical treatment of a given topic—but any advantage gained in inspiring further theory is heavily outweighed by less effective communication to the broader scientific community. Overall, equation density has a strong negative impact on citation rates and, thus, presumably impedes the wider dissemination of theoretical predictions. This finding should give pause for thought to scientists aiming to communicate theory in the most effective way. New ideas spread through a cumulative process,



**Fig. 1.** Equation-dense articles receive fewer citations from nontheoretical articles but not from other theoretical articles. The graphs show the mean ( $\pm$ SEM) number of citations by nontheoretical papers (A) and theoretical papers (B) for cited articles of differing length and number of equations per page (for the main text and appendixes combined). For illustration purposes only, the number of equations per page was binned into the ranges shown on the x axis; note that the data were not binned for the statistical analysis.







**Fig. 2.** Equations presented in the main text reduce citations from non-theoretical articles, whereas equations presented in an appendix do not. The graphs show the mean ( $\pm$ SEM) number of citations by nontheoretical papers for cited articles of differing length and number of equations per page, when those equations are presented in the main text (A) or an appendix (B). For illustration purposes only, the number of equations per page was binned into the ranges shown on the x axis; note that the data were not binned for the statistical analysis.

We examined all articles published in the three chosen journals in 1998, counting equations, inequalities, and other mathematical expressions (hereafter referred to simply as equations) in (i) the main text and (ii) any printed appendices. In 1998, online-only electronic appendices were very rare, so we ignored any that were present. We only counted equations that were presented on lines set apart from the text, but two or more such equations written on the same line were considered as separate. “In-line” equations printed fully within the text, without breaking its spacing or indentation, were not counted.

We obtained citation data for these articles from the Science Citation Index Expanded on the Thomson Reuters Web of Science in May 2011. In calculating the number of citations, we ignored self-citations by excluding any citing papers for which one or more of the author surnames matched one or more of the author surnames for the cited paper. Although we acknowledge that this criterion might generate some spurious self-citations, they are likely to be rare and so not problematic in such a large dataset. In any case, when we included self-citations, we obtained very similar results.

We downloaded the abstracts of all articles where these were available, which was for 28,068 of the 29,072 citing articles (96.5%). We then searched for the term “model\*” in the title or abstract of the citing articles (where the asterisk is a “wildcard” representing any group of characters and will therefore locate all instances of “model,” “models,” “modeled,” “modelled,” “modeling,” and “modelling”), excluding some common empirical uses (namely “model organism\*,” “model species,” “model system\*,” “model egg\*,” “model predator\*,” “experimental model\*,” “statistical model\*,” “regression model\*,” “general\* linear model\*,” and “general\* additive model\*”). We used this as a rough proxy for whether the citing paper was a theoretical one. (We felt that “theor\*” would be too broad as a search term and would identify too many general references to evolutionary theory.) The search identified 6,229 (22.2%) of the 28,068 citing articles as “theoretical,” which is likely to be an overestimate of the true proportion of theoretical studies in evolution and ecology. To check the validity of our proxy, we examined a randomly selected subset of 200 of the citing articles and recorded whether they contained a substantial mathematical component (excluding statistical analysis of empirical data). For this subset, our proxy correctly classified 84.5% of articles as theoretical or nontheoretical.

**Dataset S1** lists the cited articles and their citation data.

**Statistical Analysis.** We analyzed the citation patterns by fitting generalized linear models for count data using the statistical software package R (19). A Poisson model for the error terms was not appropriate because the data were extremely overdispersed, with a variance-to-mean ratio in excess of 50. This overdispersion is unsurprising given that successive citations of a paper are not independent events but tend to attract additional citations as the paper becomes increasingly widely read. We therefore used a negative binomial model (20), specified by the function `glm.nb` in R’s MASS library. As with Poisson regression, this function models the natural logarithm of the response variable, but unlike Poisson regression, it takes into account the degree to which the data cluster together (21), which we found to be extreme (estimated clumping parameter,  $0.663 \leq k \leq 0.942$ ; ref. 22). To check the sensitivity of our results to the model assumptions, we also fitted an equivalent set of models by using a quasi-Poisson error function (within the function `glm` in R). These models gave the same statistical conclusions and quantitatively similar estimates of the regression coefficients, so we present only the negative binomial models in the text. For each model, a plot of the residuals versus the fitted values and a normal quantile–quantile plot of the standardized residuals indicated no departure from the underlying statistical assumptions.

Rather than analyzing the effect of the absolute number of equations in an article, we divided this count by the article’s length (total number of pages) to get a measure of the density of equations. There are two reasons for doing this. First, it allows us to separate the effect of the number of equations from that of the number of pages, which are positively related ( $r_{647} = 0.257$ ,  $P < 0.001$ ). Second, it reflects our suspicion that equations may be more palatable to many biological readers if they are interspersed with plenty of explanatory text, rather than densely concentrated in a concise but heavily mathematical paper. To control for other influences on citation rate, we included the length of the article (total number of pages) and the journal of publication as additional explanatory variables. The density of equations per page and the total number of pages were both modeled as continuous variables instead of binned into categories as shown in the figures. We also included an interaction term between equation density and the total number of pages, because we suspected that heavy use of equations may be more off-putting if it extends over many pages.

**ACKNOWLEDGMENTS.** We thank Innes Cuthill, Alasdair Houston, Andy Radford, and Graeme Ruxton for discussion and two anonymous reviewers for comments. Support for this work was provided by European Research Council Advanced Grant 250209 (to Alasdair Houston).

1. Lunney D, Pearson JM, Thibault C (2003) Recent trends in the determination of nuclear masses. *Rev Mod Phys* 75:1021–1082.
2. Marcus RA (1993) Electron transfer reactions in chemistry: Theory and experiment. *Rev Mod Phys* 65:599–610.
3. Deco G, Jirsa VK, Robinson PA, Breakspear M, Friston K (2008) The dynamic brain: From spiking neurons to neural masses and cortical fields. *PLOS Comput Biol* 4:e1000092.
4. Elena SF, Froissart R (2010) New experimental and theoretical approaches towards the understanding of the emergence of viral infections. Introduction. *Philos Trans R Soc Lond B Biol Sci* 365:1867–1869.
5. Kareiva P (1989) Renewing the dialogue between theory and experiments in population ecology. *Perspectives in Ecological Theory*, eds Roughgarden J, May RM, Levin SA (Princeton Univ Press, Princeton, NJ), pp 68–88.

6. Raupach MR, et al. (2005) Model–data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications. *Glob Change Biol* 11:378–397.
7. Caswell H (1988) Theory and models in ecology: A different perspective. *Ecol Modell* 43:33–44.
8. Mock DW, Forbes LS (1992) Parent–offspring conflict: A case of arrested development. *Trends Ecol Evol* 7:409–413.
9. Brandon RN (1994) Theory and experiment in evolutionary biology. *Synthese* 99: 59–73.
10. Weiner J (1995) On the practice of ecology. *J Ecol* 83:153–158.
11. May RM (2004) Uses and abuses of mathematics in biology. *Science* 303:790–793.
12. Butlin RK, Tregenza T (2005) The way the world might be. *J Ecol Biol* 18:1205–1208.

13. Odenbaugh J (2005) Idealized, inaccurate but successful: A pragmatic approach to evaluating models in theoretical ecology. *Biol Philos* 20:231–255.
14. Kokko H (2007) *Modelling for Field Biologists and Other Interesting People* (Cambridge Univ Press, Cambridge, UK).
15. Codling EA, Dumbrell AJ (2012) Mathematical and theoretical ecology: Linking models with ecological processes. *Interface Focus* 2:144–149.
16. Levin S (2012) Towards the marriage of theory and data. *Interface Focus* 2:141–143.
17. Bialek W, Botstein D (2004) Introductory science and mathematics education for 21st-century biologists. *Science* 303:788–790.
18. Hawking S (1988) *A Brief History of Time: From the Big Bang to Black Holes* (Bantam Books, New York).
19. R Development Core Team (2011) R: A Language and Environment for Statistical Computing (R Found Stat Comput, Vienna), version 2.13.1.
20. White GC, Bennetts RE (1996) Analysis of frequency count data using the negative binomial distribution. *Ecology* 77:2549–2557.
21. Bolker BM (2008) *Ecological Models and Data in R* (Princeton Univ Press, Princeton, NJ).
22. Crawley MJ (2007) *The R Book* (John Wiley & Sons, Chichester, UK).

## No evidence that equations cause impeded communication among biologists

In their recent contribution to PNAS, Fawcett and Higginson (1) analyzed article equation density with respect to citation rate. The observed negative correlation is used to conclude that “equation density has a strong negative impact on citation rates,” which then “presumably impedes the wider dissemination of theoretical predictions” (1). This conclusion is incorrect as it mistakes correlation for causation.

Although the correlation is found to be significant, no evidence of causality is shown. It is plausible that equations cause low citation rates via impeded communication. However, equations are used specifically to communicate with precision and conciseness. Concepts expressed easily with an equation are often unwieldy when expressed via text. Thus, it is equally plausible that, if equations disappeared en masse from scientific papers, the resultant articles would be unintelligible even to experts, and thus maintain a low citation rate.

For example, the common inequality “ $P < 0.05$ ” expresses the phrase “...conditioned on the null hypothesis, the random variable  $P$  has a uniform distribution over the Borel field of the unit interval and is found to have a realized value, conditioned on the data, to be within the fifth lower percentile...” with great brevity and clarity.

Fawcett and Higginson (1) further assume that citation rates reflect communication level, and that said communication reflects understanding. However, consider the following two counterexamples. In bioinformatics, the work of Altschul et al. describing the BLAST tool (2) has 32,000 citations and counting. In statistics, the work of Fisher describing  $P$  values (3) has

transcended citation. Both these works are models of clear, minimal equation writing, yet the conclusions of each work are routinely misunderstood. BLAST  $E$  values are commonly confused with  $P$  values, and  $P$  values are commonly confused with false-positive error rates. Thus, in neither case does citation rate correspond to communication effectiveness.

In both cases, the mechanism causing impeded communication is from the reader, not the writer. Further, given the importance of these works to bioinformatics and biology in general, we reject the statement that it is too difficult or too long-term to think of strengthening the mathematical expertise of biologists. If anything, as mathematics rapidly becomes a more central tool for biology in general, we believe it is too urgent that biologists gain mathematical sophistication, and do so quickly. For biology to advance, greater mathematical literacy is an unavoidable necessity.

In his bestselling book (4), Nobel Prize-winning physicist Richard Feynman tackles the Herculean task of explaining quantum electrodynamics to the general public. His success in doing so, in his own words, comes at the price of inefficiency. Thus, although the reader will be able to understand the discipline without using mathematics, he will forever be unable to meaningfully work within it.

**Andrew D. Fernandes<sup>1</sup>**

*YouKaryote Genomics, London, ON, Canada N6K 0A1*

1. Fawcett TW, Higginson AD (2012) Heavy use of equations impedes communication among biologists. *Proc Natl Acad Sci USA* 109:11735–11739.
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
3. Fisher RA (1925) *Statistical Methods for Research Workers* (Oliver and Boyd, Edinburgh).
4. Feynman RP (1985) *QED: The Strange Theory of Light and Matter* (Princeton Univ Press, Princeton).

Author contributions: A.D.F. wrote the paper.

The author declares no conflict of interest.

<sup>1</sup>E-mail: andrew@youkaryote.com.

## Mathematical illiteracy impedes progress in biology

Fawcett and Higginson (1) have shown that citation rates of biology papers with equations in the main text are lower than those of papers without equations. They claim that this implies “heavy use of equations impedes communication among biologists” and recommend that equations be moved to appendixes to improve citation rates. We suggest that a better interpretation of their results would be “mathematical illiteracy impedes progress in biology.”

The authors base their conclusion on the assumption that a simple count of citations is a useful measure of communication and value of scientific work. However, not all citations demonstrate understanding. A citation of a paper embedded in a long list of citations to show awareness of the field is far less valuable than a citation of a paper that forms the foundation of new work. The primary goal of a scientific paper should not be to maximize its citation index but to improve readers’ understanding of physical and biological phenomena.

If ideas are expressed vaguely, it is often impossible to build on them, and mathematics (as represented by “equations” in the authors’ work) is often essential for expressing ideas precisely. An example from our own work is a model of malaria immunity that we expressed entirely in analogies (2). We have not been able to translate this into a formal description, making it es-

entially useless as a source of further scientific development. Complicated expressions do need explanatory text to ease understanding, but hiding them in appendixes only serves to obscure the readers’ understanding of the subject matter. Moving equations to an appendix means the authors are less likely to put in the required effort to explain the meaning of the equations, the reviewers less likely to check them, and the readers less likely to put in the required effort to understand the meaning. None of these consequences benefit scientific communication.

As biology focuses increasingly on complex systems and automated data gathering, mathematical and statistical literacy are more important than ever. Authors of theoretical papers in biology should include the mathematics and make the effort to explain it in an easily understandable manner. The readers should make a corresponding effort to understand that theory.

**Nakul Chitnis<sup>a,1</sup> and Thomas A. Smith<sup>b</sup>**

<sup>a</sup>*Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, 4002 Basel, Switzerland; and* <sup>b</sup>*Universität Basel, 4003 Basel, Switzerland*

1. Fawcett TW, Higginson AD (2012) Heavy use of equations impedes communication among biologists. *Proc Natl Acad Sci USA* 109:11735–11739.
2. Smith T, Felger I, Tanner M, Beck HP (1999) Premunition in *Plasmodium falciparum* infection: insights from the epidemiology of multiple infections. *Trans R Soc Trop Med Hyg* 93(suppl 1):59–64.

Author contributions: N.C. and T.A.S. wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. E-mail: nakul.chitnis@unibas.ch.

## Reply to Chitnis and Smith, Fernandes, Gibbons, and Kane: Communicating theory effectively requires more explanation, not fewer equations

Mathematical theory is an indispensable part of research in ecology and evolutionary biology. Used wisely, equations can capture the essence of fundamental biological processes with greater clarity, precision, rigor, and brevity than purely verbal arguments. In our original article (1), we showed that the citation rate of articles in ecology and evolutionary biology is negatively associated not with the number of equations used—contrary to what Gibbons (2) and Chitnis and Smith (3) claim—but with how densely packed those equations are on each page. This is a crucial difference. It is not equations that are the problem; it is equations without sufficient accompanying text to explain the assumptions and implications for a broad biological audience. We do not recommend the indiscriminate removal of equations from scientific papers.

Explaining the mathematics in sufficient detail for a broad audience can, however, require considerable space. As a pragmatic solution acknowledging the constraints many journals impose on article length, we suggested that authors might move some of their equations to an appendix. Our viewpoint is that essential equations capturing the assumptions and structure of a model should be presented in the main text, whereas less fundamental equations, such as those describing intermediate steps to solutions, should be presented in an appendix. However, we also warned that this approach requires extra care in stating the mathematical assumptions clearly in the main text (1). Chitnis and Smith (3) echo this concern, but go further by concluding that authors presenting equations in an appendix are “less likely to put in the required effort to explain the meaning.” We are not convinced this is true; in principle, authors can provide much more detailed and comprehensive explanations when writing an online appendix, for which they are freed from constraints on page space. Furthermore, by moving equations to an appendix, authors would be forced to provide a verbal description in the main text. However, we agree that the mathematical details should be easily accessible to those who want to examine them, so we welcome the suggestion by Kane (4) to build closer links between the main text and online material.

Chitnis and Smith (3) and Fernandes (5) question whether citation rate reflects how well a theory has been understood, which we agree is an important issue. It is easy to find striking counter examples, as Fernandes (5) has done, of astoundingly well-cited articles that have been misunderstood by many readers. However, our concern is with the low-citation end of the scale, where a potentially important theoretical contribution is overlooked because the mathematics are not sufficiently well

explained for a broad audience. An article with a high citation rate may have been successful for many different reasons (6, 7), but surely one that is hardly ever cited has little hope of making an impact. Furthermore, citation counts really do matter for the careers of individual scientists, who are increasingly judged by their h-index (8) and similar metrics. It would be a shame if promising scientists failed to reach their potential as a result of how they present their work.

Fernandes (5) asserts that “the mechanism causing impeded communication is from the reader, not the writer.” However, both parties have a responsibility to enhance understanding. We emphatically reject the notion that improving the communication of mathematical theory in biology and improving the mathematical education of biologists are somehow mutually exclusive goals. They are clearly complementary, and scientists who fail to recognize the importance of both solutions risk slowing progress in their field and harming their own careers.

The hypothesis put forward by Gibbons (2), that biologists are generally less interested in theoretical than empirical studies, is not supported by the data. As we originally reported (1), we find a significant negative association between equation density and citation rate even after removing papers with zero equations from the analysis. Our study found no evidence for a general aversion to theoretical papers.

Interestingly, the more complex statistical analysis of Gibbons (2) shows that the effect of equation density differs between journals, with a positive effect for articles published in *The American Naturalist*. It is important that future studies extend our analysis to other journals, as well as other years of publication and other fields. Among journals in ecology and evolutionary biology, the three we studied publish a relatively high proportion of theoretical papers (Table 1), and so are likely to be read by mathematically inclined biologists, arguably making our analysis a conservative test of our hypothesis. We expect that the effects would be even stronger for journals with a lower theoretical content.

**Tim W. Fawcett<sup>1</sup> and Andrew D. Higginson**

*School of Biological Sciences, University of Bristol, Bristol BS8 1UG, United Kingdom*

1. Fawcett TW, Higginson AD (2012) Heavy use of equations impedes communication among biologists. *Proc Natl Acad Sci USA* 109:11735–11739.
2. Gibbons JM (2012) Do not throw equations out with the theory bathwater. *Proc Natl Acad Sci USA* 109:E3054.
3. Chitnis N, Smith TA (2012) Mathematical illiteracy impedes progress in biology. *Proc Natl Acad Sci USA* 109:E3055.
4. Kane A (2012) A suggestion on improving mathematically heavy papers. *Proc Natl Acad Sci USA* 109:E3056.
5. Fernandes AD (2012) No evidence that equations cause impeded communication among biologists. *Proc Natl Acad Sci USA* 109:E3057.
6. Leimu R, Koricheva J (2005) What determines the citation frequency of ecological papers? *Trends Ecol Evol* 20:28–32.
7. Bornmann L, Daniel H-D (2008) What do citation counts measure? A review of studies on citing behaviour. *J Doc* 64:45–80.
8. Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 102:16569–16572.

Author contributions: T.W.F. and A.D.H. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. E-mail: tim.fawcett@cantab.net.



**Table 1. Proportion of articles that reported theoretical work in the top 20 journals in ecology and evolutionary biology in 2011**

Journal	5-y Impact factor	No. of papers	Theoretical proportion
<b><i>Evolution</i></b>	<b>5.6</b>	<b>304</b>	<b>0.188</b>
<i>Molecular Biology and Evolution</i>	9.9	306	0.176
<b><i>American Naturalist</i></b>	<b>4.7</b>	<b>182</b>	<b>0.154</b>
<i>Ecology Letters</i>	15.4	155	0.142
<i>Journal of Evolutionary Biology</i>	3.7	261	0.138
<i>Ecological Applications</i>	5.4	272	0.118
<b><i>Proceedings of the Royal Society of London B</i></b>	<b>5.7</b>	<b>495</b>	<b>0.109</b>
<i>Journal of Applied Ecology</i>	5.8	173	0.104
<i>Journal of Animal Ecology</i>	5.0	137	0.095
<i>Heredity</i>	4.5	185	0.081
<i>BMC Evolutionary Biology</i>	4.6	379	0.079
<i>Ecology</i>	6.0	240	0.063
<i>Behavioral Ecology</i>	3.4	201	0.060
<i>Oecologia</i>	3.9	317	0.050
<i>Global Ecology and Biogeography</i>	6.6	81	0.049
<i>Molecular Phylogenetics and Evolution</i>	4.0	257	0.047
<i>Journal of Ecology</i>	6.0	157	0.045
<i>Oikos</i>	3.6	214	0.037
<i>Ecography</i>	5.5	107	0.028
<i>Functional Ecology</i>	4.9	146	0.014

The journals used in ref. 1 are highlighted in bold. Articles were identified as theoretical if the title or abstract contained the search term "model\*", excluding some common empirical uses (as described in ref. 1).