Taylor & Francis
Taylor & Francis Group

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: https://www.tandfonline.com/loi/utas20

The Impact of Results Blind Science Publishing on Statistical Consultation and Collaboration

Joseph J. Locascio

To cite this article: Joseph J. Locascio (2019) The Impact of Results Blind Science Publishing on Statistical Consultation and Collaboration, The American Statistician, 73:sup1, 346-351, DOI: 10.1080/00031305.2018.1505658

To link to this article: https://doi.org/10.1080/00031305.2018.1505658

© 2019 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 20 Mar 2019.

Submit your article to this journal

Article views: 4252

View related articles

View Crossmark data

Citing articles: 3 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=utas20

THE AMERICAN STATISTICIAN
2019, VOL. 73, NO. S1, 346–351: Statistical Inference in the 21st Century
https://doi.org/10.1080/00031305.2018.1505658

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Check for updates

The Impact of Results Blind Science Publishing on Statistical Consultation and Collaboration

Joseph J. Locascio

Neurology Department, Harvard Medical School, Boston, MA

ABSTRACT

The author has previously proposed results blind manuscript evaluation (RBME) as a method of ameliorating often cited problems of statistical inference and scientific publication, notably publication bias, overuse/misuse of null hypothesis significance testing (NHST), and irreproducibility of reported scientific results. In RBME, manuscripts submitted to scientific journals are assessed for suitability for publication without regard to their reported results. Criteria for publication are based exclusively on the substantive importance of the research question addressed in the study, conveyed in the Introduction section of the manuscript, and the quality of the methodology, as reported in the Methods section. Practically, this policy is implemented by a two stage process whereby the editor initially distributes only the Introduction and Methods sections of a submitted manuscript to reviewers and a provisional decision regarding acceptance is made, followed by a second stage in which the complete manuscript is distributed for review but only if the decision of the first stage is for acceptance. The present paper expands upon this recommendation by addressing implications of this proposed policy with respect to statistical consultation and collaboration in research. It is suggested that under RBME, statisticians will become more integrated into research endeavors and called upon sooner for their input.

ARTICLE HISTORY
Received February 2018
Revised April 2018

KEYWORDS
Publishing; Publication bias; Results-Blind; Significance testing

1. Introduction

In earlier articles (Locascio 1999, 2011, 2015, 2017a, 2017b), I suggested and defended an approach to evaluating manuscripts submitted to scientific journals (especially in behavioral and medical sciences), I referred to as "Results Blind Manuscript Evaluation" (RBME). In the present article, I would like to expand on implications of RBME with respect to statistical inference and methodology, focusing mostly on the role of statistical consultants and collaborators in research and in preparing reports thereof. I will first briefly summarize what RBME is and preliminary to that, the problems it is intended to ameliorate.

2. Problems in Scientific Publication and Statistical Inference

Part of the impetus for this special issue of *The American Statistician* is the intensified controversies that have recently arisen regarding problems in statistical inference and science publishing, especially the interrelated concerns of publication bias, the overuse and misuse of null hypothesis significance testing (NHST), and claimed irreproducibility of published scientific findings.

Publication bias, as meant here, is the widely recognized and documented practice in which, all else being equal, null findings of even well conducted studies are not as likely as positive findings of studies to be reported and published in scientific journals, either because manuscripts reporting such are not as likely to be accepted by journals (Fanelli 2010; Van Assen et al. 2014) or because of an inhibition among investigators to even write up or submit such reports to journals because of this assumed bias at the journals (the "file drawer" problem discussed by Rosenthal 1979). As the result of publication bias, to the degree that true effects are negligible and this bias is operative at the author submission and/or journal acceptance stage, the literature becomes contaminated with a misleadingly higher proportion of false positive findings than would occur by chance alone if there were no such bias. Further, awareness of publication bias can create a temptation among investigators to deliberately doctor study results to show positive findings in order to get the publications seen as vital to career advancement. And what may be far more commonly practiced are the many ways that unintentional and unconscious processes and choices may be at play in how researchers collect and analyze their data in order to obtain the "statistically significant" findings they feel they need for publication (e.g., "p hacking" and "opportunistic biases"; DeCoster and Sparks 2015). The implications of publication bias in meta-analysis especially has often been cited, and a number of methods that attempt to assess and correct for it have been proposed (e.g., "funnel plots"; Begg and Mazumdar 1994; Egger et al. 1997; Light and Pillemer 1984), though a better solution would seem to be to avoid the bias at its source.

"Statistically significant" results as obtained by NHST have to a great degree become collectively accepted in behavioral and medical sciences as the gold standard indicator of what

CONTACT Joseph J. Locascio, Ph.D. *JLocascio@partners.org* Harvard Medical School, and Memory and Movement Disorders Units, Massachusetts Alzheimer's Disease Research Center, Neurology Dept., Massachusetts General Hospital (MGH), 50 Staniford St., Office 560-H, Boston, MA 02114.

© 2019 The Author. Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

constitutes the positive findings that journals are biased toward accepting. However, serious and valid concerns about the overuse and misuse of NHST have been cited by statisticians for decades. (There is no space to go into these criticisms again here, but you can see, e.g., Cohen 1994, Hunter 1997, and Seife 2015, for a sampling of them.) More highly regarded alternatives to NHST have been widely recommended (e.g., confidence intervals, effect size estimates, Bayesian statistics, graphical analysis, etc.) though none of these has come close to replacing NHST as the workhorse of statistical inference.

There has also been recent concern about a recognized and documented frequent failure to replicate findings in scientific reports ("the crisis of irreproducibility"; see, e.g., Open Science Collaboration 2015). The reasons for these failures have been heavily debated with publication bias being commonly alleged to be one of the factors responsible, which includes biases at the investigator level whereby conscious and unconscious stratagems are employed to get "significant" results. It has also been claimed that, specifically, the bias of publishing only results with NHST p-values below the widely accepted 0.05 or 0.01 cutoffs has made much failure to replicate inevitable given statistical "regression to the mean" (Trafimow and Earp 2017).

3. Results blind Manuscript Evaluation (RBME)

As a method of mitigating the problems of statistical inference and scientific publication noted above, I have suggested that behavioral science and medical journals consider instituting a policy of results blind manuscript evaluation (RBME) of the suitability of submitted manuscripts for publication. For a detailed description of RBME see Locascio (2017a, 2017b) and also Locascio (1999, 2011, 2015). By RBME, I mean that in deciding whether to publish a manuscript, journal reviewers and editors would give no weight to reported results in making this judgment. Reports would be judged *exclusively* on the basis of the perceived: (1) *importance* of the substantive issues addressed, which is generally communicated in the "Introduction" section of the manuscript, and (2) soundness of the *methodology* employed, which would include among other things, appropriate: materials, measurement methods with e.g., up-front reliability/validity estimates, subject selection, study layout/experimental design, proposed data analysis techniques, and sample size, all conveyed in the "Methods" section. Regarding appropriate data analysis methodology, many proposed methods recommended as alternatives to NHST could potentially be employed, instead of or in addition to NHST. Use of NHST, if judged applicable, and correctly used and interpreted, could possibly be permissible as providing one supplementary piece of evidence that might be pertinent to research questions, at the discretion of reviewers. However, importantly, obtaining an NHST p-value less than an arbitrary cutoff would no longer be a necessary precondition for publication – it can't be because neither p-values *nor any other result* would be considered in the decision to publish. This RBME editorial policy would be made explicit and prominently stated as a part of journals' guidelines for authors so that investigators considering manuscript submission would be fully aware of it and not mistakenly censor or inhibit their own submissions because null findings were obtained, and they think lack of positive or statistically significant findings precludes publication or lessens the likelihood of it.

Note that I am recommending results blind *evaluation* of manuscripts for suitability for publication, not a complete results blind *review*. Reported results sometimes indicate some aspects of methodology and data analysis techniques which could not be fully elaborated in the Methods/Data Analysis sections and reflect good or poor methodology. Further, Results and Discussion sections certainly have to be edited and reviewed prior to the manuscript going to print, for presentation as well as for substantive reasons. Data analysis in the Results section has to be checked for execution, and correct interpretation of results in the Discussion section has to be assessed as well as a statement of limitations of the study. By "results blind evaluation", I merely mean that the nature of the observed findings of a study, or whether or not the study found any effects at all, should be given little or no weight and have no direct bearing on whether the manuscript is to be *accepted for publication*.

As a method of practically implementing RBME, I recommend a *two stage* procedure whereby authors would submit a manuscript in its entirety to a journal, just as they normally do now, seamlessly without change. But, upon receipt of it, the editor would distribute it to reviewers after having stripped it of everything except the Introduction and Methods sections (and any appendices or supplementary information relevant to methods), in order to obtain a "first stage *evaluation*" of suitability for publication. If the decision of reviewers at this stage is to reject, the authors would be informed that their manuscript was rejected and why. However, manuscripts passing this first stage would then have their Abstract, Results and Discussion sections re-attached by the editor and submitted to reviewers for a "second stage *review*". A manuscript surviving the first stage as acceptable or judged as requiring no more than minor modifications/revisions would be considered "conditionally accepted" for publication, that is, conditional on the absence of any new evidence of flawed methodology discovered in the second stage sufficiently grievous to override the first stage decision. Note the second stage review merely serves a disconfirmatory or veto function, but where still no weight is given to what the reported results are per se, in the decision to publish. After completion of the second stage, a formal decision is finalized for acceptance, acceptance with revision, or rejection (in the unlikely event rejection is decided at the second stage), and if not rejected, recommendations for minor revision, editing, wording and cosmetic alterations, etc. would be suggested or made as pertinent. Such a process would arguably involve about the same amount of labor for reviewers as is currently the case, and possibly less, given that reviewers will sometimes only have to review half a manuscript if the first stage evaluation results in rejection. If it is accepted during the first stage, and the same reviewers who conducted the first stage evaluation, perform the second step review, there is no more additional workload than had they reviewed the entire manuscript at once from the onset.

RBME would I believe ameliorate (though not entirely solve) all three of the problems with scientific publication cited above, that is, publication bias, misuse of NHST, and irreproducibility. That it would reduce publication bias is essentially self-evident. There would be no results based bias at the journal because reviewers cannot decide acceptance based on results if

they do not know what the results were. There would be little or no self-censoring bias among investigators because, being fully informed of the RBME journal policy, they would presumably have no inhibition in reporting the null findings of a well conducted study, knowing that the absence of positive findings per se will in no way reduce their chances for publication and getting just credit and recognition for their hard work.

Although RBME is somewhat orthogonal to many of the claimed problems with NHST, indirectly, many of these endlessly criticized and lamented issues will be mitigated or made nonissues. Although NHST need not necessarily be explicitly banned by anyone, authors would feel no compulsion to employ that particular methodology, or any other for that matter, beyond what they deem as most relevant, fitting, appropriate, methodologically sound, justifiable, and defensible for their study. Controversy will no doubt continue over whether NHST has some limited utility in science given proper circumstances of relevancy, and correct application, implementation and interpretation; however the widely criticized overuse, misuse and over-emphasis of NHST p-value cut-offs as the predominant gate-keeper of communication of scientific results will end. Although cautions will still be cited (and should be) regarding problems with misuse of NHST, the *practical imperative* of NHST, that is, the widely perceived necessity of employing NHST and getting $p < 0.05$ for publication, which implicitly fuels and exacerbates many of these problems, will be over. Researchers will no longer feel any particular need to get $p < 0.05$ or $p < 0.01$ or any other p-value. However, if they are using NHST, it will now become important for them to defend that they are using it correctly and, unlike before, justify why they are using it at all.

The problem of irreproducibility of results will also presumably be mitigated as well because there will likely be a reduction in published false positives given that positivity of findings per se will no longer consciously or unconsciously influence any decision to report study results or publish them. Null findings judged to be of equal validity to any positive results because they were based on equally sound methodology applied to equally important research questions will have the same chance of being published as the positive results. Thus, the scientific literature will convey a representative, balanced sample of findings containing a proportion of negative and positive findings duly reflective of what is actually likely to be true. Any future, well conducted replication work would likely reproduce what has already been generally found by previously published well conducted studies, and contribute to accumulating supportive evidence in further reinforcing truthful scientific theories and refuting false ones.

RBME is not claimed to be panacea for all problems of statistical inference and publication. Some forms of bias in publishing will still remain, e.g., bias by reviewers and editors as to what is an "important" research question and good methodology, and notably biases that do not occur at the point of journal *reviewer evaluation*. For example, bias at point of author *submission* can still occur to the extent that the decision to submit is based on study results being supportive of personally favored and espoused theories, or their being in line with monetary interests, or because of other nonscientific incentives. A pharmaceutical company may only submit reports of studies whose results are seen as favorable to their product. Or investigators may only submit reports of studies that ostensibly confirm their pet views. Pre-registering studies help avoid some of these problems, but of course they involve more elaborate procedures than RBME which is more practical as a widespread journal practice. Further, publication bias toward positive results can still be exercised by the journal *editor* with the pre-decision to decline to forward a manuscript for further review. However, in my experience, a "preemptive rejection" such as this is usually due to a manuscript not meeting author guideline requirements or because the substantive content of the paper is judged to clearly not be a good fit for the given journal. An editor who has already made the decision to adopt an RBME policy for the journal, would not seem likely to then consciously turn down manuscripts based purely on results in direct contradiction of the policy. And finally, bias contingent on observed results is still possible to some extent by reviewers at *stage 2* of RBME. However, under RBME this can happen only *after* a reviewer has already committed to a conditional acceptance for publication, a decision that could not have been based on results because they were not seen. Results can influence the reviewer's decision regarding publication only in a *negative* manner, that is, to subsequently want to reject that for which one has already formally stated an acceptance. The results wouldn't be reviewed without the initial acceptance. Bias based on results cannot work in an *affirmative* manner, that is, results cannot influence the reviewer to accept a report based largely on results of a study that was viewed as otherwise not worthy of publication. Furthermore, by the rules of RBME, a second-stage rejection will require the reviewer to justify to the editor that his/her change of mind is based on specifically *methodological* problems discovered in the Results or Discussion sections, problems serious enough to override the reviewer's original decision and not being amenable to revision.

There are other limitations of RBME, many of which are stated and addressed in detail in Locascio (2017a, b). For example, some serious methodological flaws may only be evident in the Results and Discussion sections of manuscripts. However, rejection at the second stage of RBME, noted above, would hopefully catch those. Further, not every contingency can be anticipated. For example, it's conceivable that a compelling result relevant to a medical emergency might supersede minor methodological considerations. I do not advocate that RBME be enforced rigidly without occasionally considering, on a case-by-case basis, possible exceptional circumstances, and there is nothing to preclude authors from arguing in a cover letter to the editor such extenuating circumstances. Lastly, one of the positive consequences of RBME would be to accentuate the need for a sample size adequate to address a study's goals, but this may be seen as unfairly favoring large universities and medical centers with access to large subject pools. In a sense, this is unfortunate, but it would provide an incentive for researchers to develop resources for greater participant access, and besides, if we have to be unfair in some way, is being unfair in favoring better studies so bad? I can think of worse ways of being "unfair".

Naturally, because of these and possibly other unforeseen practical and theoretical problems with RBME, it is recommended that RBME be introduced in a gradual, experimental manner with on-the-fly modifications as might be warranted.

4. How RBME Would Impact Statistical Consultation and Collaboration

Researchers want to do good research and they want to get published. Unfortunately, our existing publication systems sometimes put these two goals at odds. With RBME, I believe they would become better aligned. Claimed findings being reported would no longer per se have any bearing on publication and dissemination across the scientific community. This criterion for publication would instead be replaced by whether what is being reported is likely to be true, and the latter depends mostly on sound methodology, described largely in the Methods section of manuscripts. (The judged substantive importance of research questions addressed in a study, conveyed for the most part in the Introduction section, which is presumably a factor in publication now, would also continue to be so under RBME.) The switch of emphasis from *what* is claimed to have been found to *how* it was found, and from what was claimed as the answer to a question, to what the question was and how accurately it appears to have been answered in determining potential for publication, would have a profound impact on the role of statistical analysis, methodology, and statistical and methodological consultation/collaboration in research. At least the first provisional decision regarding acceptance for publication can no longer be based on the results of a study because they were not looked at in making the decision. The importance of the described methodology would, however, become magnified. It would now be critical that the study's methods are sound, thoroughly described and well justified, including the proposed data analysis techniques to be employed, and justification for the sample size. This means statisticians and methodologists will become increasingly needed as a very practical necessity, even if, regrettably, not for more elevated reasons.

4.1. Statisticians Would Become More Involved in Research Studies

I am a member of a university based biostatistics consulting group, which provides statistical consultation and analysis assistance for studies conducted across research facilities throughout a large metropolitan area. In this role, it is more often the case than not, that I'm not merely asked to answer specific questions on an otherwise fully developed research study. Rather, oftentimes, the investigators have a somewhat vague idea for a study they wish to conduct, or are articulating it poorly, and I find myself trying to solve an equation in two unknowns. First, I need to clarify for myself, and for the investigators, what exactly the research question and hypotheses are, and then having done that, translate this into the specifics of their study situation, and, finally, recommend statistical tools that seem to be a good fit for addressing those questions and hypotheses in light of anticipated or collected data. Before long, I feel myself, often with mutual consent, being pulled into the guts of the study, not merely serving on the periphery to answer a few specific technical statistical questions. Without my having asked, I'm sometimes invited to be a co-author on an anticipated manuscript or even to get on board a grant.

If I assume the researchers with whom I consult are somewhat representative of many others I've never met in the field, it's clear to me that many investigators in medical and behavioral sciences could benefit from statistical collaboration, beyond just spot consultation and analysis. Under known RBME journal policies, these investigators may feel a more imperative need to seek such in-depth statistical and methodological assistance. Statisticians might be more frequently brought on board in a study as an integral part of the research team, not just an outside "technician". This change would be to everyone's benefit.

Under RBME, researchers will at least appreciate the practical benefit of integrating a statistician into their research project. As such, I predict that the statistician will often need to struggle with the investigators to get a clear idea of what the basic research questions are and help articulate them well. In the past, for this purpose, I've often found it helpful to sketch rough causal models as path analysis diagrams (or if you prefer, graphs of structural equation models, SEM, directed acyclic graphs, DAG, etc.). These models force everyone to be very explicit, specific and thorough on just what are the underlying causal mechanisms that are at the root of their hypotheses or research questions, what are the relevant exogenous and endogenous variables, and what covariates might be important and why. Once there is consensus on underlying models to be investigated, it would be primarily the statistician's task to select from his/her hopefully diverse toolbox of statistical techniques, those that seem most congruent with the study's analytic requirements.

As a statistician myself, predicting a greater need for statisticians under RBME may seem self-serving, but I truly believe that such increased involvement will ultimately be of benefit to the research community in general and the advancement of science.

4.2. Statisticians Would Become Involved in Studies Earlier

In my role as statistical consultant for research, I'm often asked for assistance very late in the game, so to speak. Someone has designed a study and collected a lot of data, and now wants to talk about statistical analysis. Or else, revisions to a manuscript submitted to a journal have been requested by reviewers including or especially regarding data analysis, and the investigators need the advice of a statistician and/or help in analysis. In these cases, oftentimes I feel what the study really needs is a new design and new data, more fitting to the purpose of the study, once I discover what that is, but it's too late for that. I try to be sensitive and do the best I can by recommending and conducting ad hoc statistical patch-ups, making clear the limitations of these methods, but these situations can be frustrating. Under RBME, investigators will know in a very practical sense that they must seek needed statistical help *early* in the research process, i.e., as they begin to design a study, if possible before data are collected, or at least before they submit a manuscript for publication, because under RBME they know the proposed research design and data analysis methods described in their report will heavily influence the decision to publish or not. The obtained results will not be looked at in the initial decision and have little or no influence on it. Being informed by an RBME rejection that there are fundamental problems in a study's design, may unfortunately sometimes be too late for any corrective action to be taken for that particular study, but the situation will at

least be a learning experience for future reference and promote greater and more timely care in subsequent work.

The statistician will also more likely become involved in a formal justification of sample size at the beginning of a study. This justification would not necessarily entail conventional power analysis, which assumes NHST, but could also involve estimating the sample size needed for a given pre-specified precision in estimating effects. There are also other formal methods of deriving sample sizes that avoid standard power analysis and NHST (see, e.g., Trafimow and MacDonald 2017). To whatever degree possible, the statistician should also monitor, oversee, and help with preparing, proofing, checking (with graphs, programming checks, etc.), and cleaning the data set, right from the beginning, instead of finding out accidentally, much later after analyses are run, that there are obvious data level problems.

In fairness to principal investigators who are not statisticians, the potential of RBME to require statistical advice earlier in the study development, may not be seen as a boon to those investigators, especially if funds for statistical consulting are not available until later in the process. However, this may create a greater impetus for investigators to sharpen their knowledge of research design and data analysis methods, including power analysis. In this internet age, there are a proliferation of good online tutorials and videos for self-education in these areas, that would hopefully provide a groundwork until formal statistical assistance can be procured.

4.3. The Statistical Analysis Section of Manuscripts Would Expand

Under RBME, authors of research manuscripts will be fully aware that the decision to publish their paper will be heavily influenced by their reported data analysis strategy. As a result, I believe it will be quite likely that the portion of the Methods section that describes statistical analysis methods will become more expanded and/or, if space limitations are a problem, more detail will be provided in an associated appendix or supplementary materials to that end (to also be reviewed in stage one of RBME). More attention will be paid by investigators to improving the quality of the statistical methodology and reporting it thoroughly. This will mean that investigators will first need to be clear, explicit, and specific about what their research questions and/or hypotheses are, and then in consultation with statisticians and methodologists, make sure they choose a research design appropriate to addressing those questions and hypotheses, which in turn should partly dictate the kinds of data analyses they need. A statistical analysis section might be further subdivided into subsections describing (a) research questions and hypotheses, (b) research design, (c) sample size justification, (d) data analysis techniques, and (e) assumptions of analysis methods and checks to verify they were met. On a related point, the section describing measurement tools employed in the study, will have to give special attention to establishing the reliability and validity of measures, preferably by using data from the current study, instead of or in addition to assessments reported in the literature. Note that results of biometric/psychometric analyses, as well as other preliminary "results", e.g., tests to show covariates were well balanced across study groups, manipulation checks, treatment implementation checks, etc. might more appropriately go in the Methods section to be reviewed in the first stage of RBME. Such information bears on the quality of the methodology, not the answer to the research questions. An expanded data analysis section such as described of course presupposes more care in data analysis and active consultation/collaboration by statisticians and research methodologists.

Authors submitting a manuscript to a journal should be fully aware of a journal's RBME policy, given that it should be explicitly stated and emphasized in the journal's guidelines for authors. Thus, authors should be motivated to explain more of the details of their data analysis methods, and other design issues, e.g., how attrition or nonresponse is to be handled, in the Methods-Data Analysis section, not just in the Results. This should lessen the likelihood of reviewers discovering methodological problems for the first time at Stage 2 of RBME.

4.4. Statistical Knowledge and the Role of Statisticians as Informal Educators Would Become more Important

Under RBME, I believe technical issues such as controversies regarding the use of Bayesian versus frequentist approaches to statistical inference will no longer be debated only among statisticians and/or in exclusively statistical journals whose contents seem esoteric to many other researchers. Investigators will of necessity have to at least familiarize themselves with core issues of current statistical controversies and form some educated opinions. A statistician can help them do that in a role of informal educator and mentor. As a statistical consultant, I try to informally teach substantive investigators the basics of what I'm recommending or doing as an analyst. I never "dump" esoteric printouts on people with whom I consult, without offering and recommending to them that I walk them through the statistical output, at least the first time through for new output they haven't seen before. I also try to provide information-rich graphical illustrations when possible that show in a more salient and intuitive way an otherwise very technical result (e.g., a graph illustrating a model predicted interaction effect or a curved quadratic function of predicted values, rather than just the corresponding regression coefficients). Many of these graphs I produce are for internal use only and are never published or seen by outside audiences. Further, I write descriptions for investigators of statistical methods I'm using and results I've obtained that are understandable to nonstatisticians, and the nonstatistical audience who may ultimately also read them, without omitting important statistical details. Aside from informal education, with RBME, statisticians may also become in more demand as formal teachers in academic settings.

4.5. Statisticians would become more Needed by Journals to Review Manuscripts

Since under RBME, the initial decision regarding acceptance for publication will be heavily based on the statistical analysis methods proposed in the Methods section (among other things included under Methods and the Introduction), "statistical reviews" of manuscripts at journals will become more often a necessity. Statistical considerations must be addressed such as justification of sample size, whether the statistical methods are a good fit for the study objective, the design and data, whether

statistical assumptions are going to be checked, and how, justification of reliability and validity of measures, justification of use of NHST and/or a Bayesian approach, etc. I am a member of the statistical review boards for two academic journals and I often find myself requesting large scale makeovers of the design and data analysis of a study, suggesting to me that there was no involvement by a statistician in the study. Under RBME, authors will know that reviewers with advanced statistical training will likely be examining their submitted manuscript and will have a big say in deciding whether to publish it or not.

5. Conclusion

In addition to other cited benefits of RBME in mitigating publication bias, problems with NHST, and irreproducibility of research findings (Locascio 2017a, 2017b), I believe RBME will provide an inducement to better quality of scientific research methodology and data analysis, greater self-education of researchers in statistics and methodology, and/or increased collaboration between principal investigators and statisticians. The emphasis of research reports will shift to a more equitable balance between reporting what was found and establishing that what is reported is likely to be true. And this shift will necessitate greater input of statisticians in research projects. As technology advances, specialization tends to do so also, and it soon may become commonplace to include a statistical and methodological specialist on research teams for virtually all studies. Most researchers cannot be expected to keep abreast of all the rapid and complex advances in their substantive field while simultaneously staying up to date on the latest equally rapid and complex advances in data analysis methods. RBME should fit right into these general trends.

Acknowledgements

Thanks to Dr. Ron Wasserstein for suggesting the idea for this paper, i.e., implications of RBME on statistical consultation.

ORCID

Joseph J. Locascio http://orcid.org/0000-0003-3439-1209

References

Begg, C. B., and Mazumdar, M. (1994), "Operating Characteristics of a Rank Correlation Test for Publication Bias," *Biometrics*, 50, 1088–1101. [346]

Cohen, J. (1994), "The World is Round ($p < .05$)," *American Psychologist*, 49, 997–1003. [347]

DeCoster, J., Sparks, E. A., Sparks, J. C., Sparks, G. G., and Sparks, C.W. (2015), "Opportunistic Biases: Their Origins, Effects, and an Integrated Solution," *American Psychologist*, 70, 499–514. [346]

Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997), "Bias in Meta-Analysis Detected by a Simple Graphical Test," *British Medical Journal*, 315, 629–634. [346]

Fanelli D. (2010), "Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data," *PLoS ONE*, 5, e10271. doi:10.1371/journal.pone.0010271 [346]

Hunter, J. E. (1997), "Needed: A Ban on The Significance Test," *Psychological Science*, 8, 3–7. [347]

Light, R. J., and Pillemer, D. B. (1984). *Summing Up: The Science of Reviewing Research*, Cambridge, MA: Harvard University Press. [346]

Locascio, J. J. (1999), "Significance Tests and 'Results-Blindness,' " *American Psychological Association (APA) Monitor*, May, 1999, pg. 11. [346,347]

——— (2011), "A Statistical Suggestion," *Scientific American Magazine*, Letters, 305, 8–10. [346,347]

——— (2015), "Psychology and Statistics: Readers Respond to the BASP Ban on *p*-Values," *Significance Magazine*, 12, 35–36. [346,347]

——— (2017a), "Results Blind Science Publishing," *Basic and Applied Social Psychology*, 39, 239–246. [346,347,348,351]

——— (2017b), "Rejoinder to Responses to 'Results Blind Publishing'," *Basic and Applied Social Psychology*, 39, 258–261. [346,347,348,351]

Open Science Collaboration (2015), "Estimating the Reproducibility of Psychological Science," *Science*, 349, 943. [347]

Rosenthal, R. (1979), "The 'File Drawer Problem' and Tolerance for Null Results," *Psychological Bulletin*, 86, 638–641. [346]

Seife, C. (2015), "Statistical Significance," in *This Idea Must Die: Scientific Theories That Are Blocking Progress*, ed. J. Brockman, New York: HarperCollins Perennial, pp. 519–522. [347]

Trafimow, D., and Earp, B.D. (2017), "Null Hypothesis Significance Testing and Type I Error: The Domain Problem," *New Ideas in Psychology*, 45, 19–27. [347]

Trafimow, D., and MacDonald, J. A. (2017), "Performing Inferential Statistics Prior to Data Collection," *Educational and Psychological Measurement*, 77, 204–219. [350]

Van Assen M.A.L.M., Van Aert R.C.M, Nuijten M.B., and Wicherts J.M. (2014), "Why publishing everything is more effective than selective publishing of statistically significant results," *PLoS ONE*, 9, e84896. [346]