Published in final edited form as:

Basic Appl Soc Psych. 2017 ; 39(5): 258–261. doi:10.1080/01973533.2017.1356305.

Rejoinder to Responses to “Results-Blind Publishing”

Joseph J. Locascio

Harvard Medical School

Abstract

I thank the accomplished investigators who offered their commentary on my article advocating results-blind manuscript evaluation (RBME). I very much appreciate their careful review of my proposals, their thoughtful comments, and thought-provoking constructive criticism. I respond to each individually.

Response to Dr. Hyman’s critique

I don’t think even a large-scale adoption by journals of an RBME policy would necessitate bringing into great question the validity of articles already published, as Dr. Hyman suggests. First, the literature is *already* under question because of publication bias, problems with null hypothesis significance testing (NHST), findings of poor reproducibility, and so on, and will be no more or less so if proposed methods to ameliorate these problems are implemented. During a recent presentation of mine of problems with NHST, a colleague asked me facetiously whether my criticisms of NHST meant that all previous published results of our research unit will now have to be retracted, which would of course be a nonsensically drastic action. But a more refined version of a statistical or methodological technique should not be withheld from use for fear that results obtained with the older method, though no doubt still largely valid, might be called into question.

Dr. Hyman (2017/this issue) indicates that “publishing an initial work with null results would discredit the remaining stream, and publishing subsequent work with null results would discredit the previous stream” (p. 248). Publishing one or two studies with null results would generally not discredit a large body of other published positive findings addressing the same question, but publishing a high proportion of null findings from methodologically sound studies would, yes, lead to some questioning of the positive results, and it *should*. In my opinion, Dr. Hyman is citing an asset of RBME, not a liability.

A possible nontrivial increase in operational costs due to RBME is cited by Dr. Hyman in that he feels editors would have to do a previously unnecessary up-front review and possible awkward redacting of the Introduction and Method sections of manuscripts to make sure they do not reveal results. In my experience, manuscripts rarely reveal results in these sections. Further, I indicated in my article that journal submission guidelines for authors

CONTACT Joseph J. Locascio jlocascio@partners.org, Harvard Medical School, Neurology, Boston, MA 02115.

ORCID

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

should explicitly and emphatically state that the results of a study will not have any bearing on the decision to publish the report of it and to ensure that authors should not reveal or clearly imply results in the title, Introduction, or Methods sections of the manuscript. If after a brief reading the editor sees any violations of these guidelines, aside from what can be immediately corrected by trivial editing, she should inform the authors that their currently submitted draft is declined for not adhering to guidelines and suggest resubmission after the authors correct the problem, just as editors do now for noncompliance with journal requirements. There is no need to waste the time and effort of an editor in inappropriately performing what should be the authors' task of ensuring that papers conform to submission guidelines.

Dr. Hyman cites ways in which fraudulent methods can be used to "game" and get around the intent of RBME, to which I can only reply that there is no system that is perfect or immune from all possible abuse, and it would be the task of editors and reviewers to address these possible problems as best they can, raise questions with authors as warranted, and reject publication if necessary, as is the case now. Given that under RBME, likelihood of publication now depends on presented methodology and not reported results, unscrupulous authors might try to deceive reviewers in various ways about the quality of their study's methodology instead of falsifying the results as in the past. But, even in this case, I would think being deceptive about methodology will be a more difficult trick than reporting unconsciously biased or deliberately falsified results. I make no claim that RBME will solve all problems of science publishing, but unattainable perfection is I believe not a sufficient reason for not implementing what I feel would still be largely a net improvement over current practices.

Dr. Hyman's suggestion that RBME be tested through small-scale experimentation before any kind of large-scale implementation is a good one. Unforeseen negative consequences are, of course, always a possibility. One can never rule out that there won't be some practical feasibility problems or overlooked negative consequences when adopting a new practice amid real-life complexities however unimpeachable it appeared in theory. It would be wise to first test RBME on a trial or pilot basis and/or implement it gradually with ad hoc revisions if needed.

Response to Dr. Marks's critique

Dr. Marks (2017/in press) argues that there is a slow trend toward moving away from basing acceptability for publication on results of NHST, suggesting that RBME may therefore be unnecessary for dealing with problems related to use of NHST. However, even if this is true, the problem of publication bias remains if publication is affected by the reported effect size of results, whether p values are a primary or supplementary index of these effects, or not involved at all. My argument is that the decision to publish should not be influenced by reported results and effect sizes, *whatever* the operational definition of the effect size, be it NHST or anything else. The decision to publish the report of a study should be based solely on the importance of questions asked by the study and the soundness of the methodology employed to answer them, not reported results. Publication should be based on what the *question* was and *how* it was claimed to have been answered, and not be influenced by what

the *answer* was claimed to be. Most of us acknowledge the value of double-blind research study designs; maybe it's time for blind research publishing also.

Dr. Marks also makes the point, with which I agree, that correctness of interpretation of results and discussion of implications of results, addressed in Results and Discussion sections of manuscripts, are very important, and manuscripts that reviewers feel substantially fail in these regards should not be published, unless these perceived shortcomings can be rectified. Note the issue here is validity of *commentary* on the results, not what the results were per se. However, I provide a mechanism for dealing with these problems via my proposed second-stage review. Notwithstanding that, I still feel the sine qua non for publishability should be sound methodology. If the study fails in that, then there is no need to go any further because there are no demonstrably valid results that need to be interpreted or whose implications need be discussed. I do indicate that the second-stage review may mandate revisions considered imperative and do leave open the option that a manuscript conditionally accepted during the first-stage evaluation might have to be rejected at the second stage if problems in interpretation and discussion of results, or execution of data analysis, cannot be resolved. Perhaps Dr. Marks was correct in indicating my reference to such required revisions as "minor" was misleading in terms of the potential importance of these issues.

Regarding Dr. Marks's claim that RBME would require more work on the part of editors and reviewers, I don't think editors will need to "scour," as he says, submitted manuscripts. As I indicated in my response to Dr. Hyman's similar point, journal guidelines will prohibit authors from revealing results in the title, Introduction, and Methods sections of the paper. The editor will merely need to briefly read the Introduction and Methods sections to make sure results are not revealed there before distributing them to reviewers. If the editors miss something, the reviewers will surely catch it and save themselves further work by informing the editor that stated submission guidelines of not reporting results in the Introduction or Methods were violated and the manuscript must be declined because of that. Dr. Marks's other points regarding increased work for reviewers may in some cases be valid, but I still feel, the lack of any need to even look at the Results and Discussion sections of manuscripts rejected at the first stage will reduce their work-load overall.

Certainly Dr. Marks makes a valid point that researchers are often disappointed when positive results of a study are not obtained that would confirm their pet theories and hypotheses. Because of that, there may still be an inclination of researchers to engage in various consciously or unconsciously biased practices to produce positive results. Thus, an RBME policy may not entirely remove publication bias in favor of positive results, at the author stage. However, with RBME, researchers who obtain undesired negative results will at least have the compensation of knowing that the hard labor they expended in meticulously running a methodologically well designed study will still have just as much potential in paying off in a publication that may advance their careers as if findings had been positive. Balancing that against possible rejection because of detection by journal reviewers of unsound methodology designed to obtain positive results, to say nothing of the risk of publicly exposed dishonesty, will hopefully greatly diminish publication bias at *both* the author and journal levels.

As Dr. Marks argues, even under RBME, NHST will no doubt continue to be widely used by the research community, perhaps for a long time, for various reasons he gives. However, now a valid *justification* in the Methods/Data Analysis section, of *why* the authors are using NHST will be seen as more conducive to getting published than obtaining any result that meets an arbitrary NHST p value cutoff. I think that is an improvement. Actually, the issues meant to be addressed by RBME are somewhat orthogonal to the controversy regarding the validity of NHST, and I'm not claiming that RBME will resolve those arguments. However, as an indirect effect, I feel RBME will ameliorate some of the widely recognized problems with NHST or make them nonissues. It will tend to end NHST to whatever degree this practice becomes widely regarded by the research community as methodologically problematic, but it will be up to the research community to address the validity of NHST as a separate issue. Perhaps, in the long run, NHST will come to be regarded as just one among many statistical inference techniques, one that may have some supplemental value in certain contexts, or no value at all as some argue (Trafimow, 2014; Trafimow & Earp, 2017; Trafimow & Marks, 2015). Importantly, though, RBME will certainly end the long-standing practice of using the $p < .05$ cutoff as the virtual gatekeeper of what is published and accepted as scientific truth, probably the most universally denounced problem with NHST, because everyone will know that obtaining a $p < .05$, or any other result for that matter, will no longer have any bearing whatever on whether the study's report gets published. With RBME, investigators will no longer feel any need to use NHST beyond their judgment of its methodological appropriateness for their study, a judgment subject to review, incidentally, by the journal to which they submit their report, which is as it should be.

Response to Dr. Grice's critique

Dr. Grice's (2017/in press) criticisms of RBME center largely around issues of problems with NHST. He suggests that RBME might "not possess the necessary teeth" to induce an abandonment of the large-scale use of NHST and the mind-set associated with NHST methodology (p. 255). He suggests that banning NHST outright from journals, as BASP has done, may be the only effective remedy to this problem, although he also implies that there may possibly be some potentially valid, limited uses for NHST or methods similar to it (e.g., with regard to randomization tests).

My response to Dr. Grice's point of view would partly reiterate what I said in response to Dr. Marks, that is, my objective in proposing RBME was not necessarily to end NHST or to directly or immediately provide a solution to problems associated with its widespread use. My immediate, primary concern was publication bias, with amelioration of some NHST problems as a potential side benefit and beneficial by-product. The question of NHST's validity would still need to be argued, and warnings issued against it as warranted, as a separate issue, orthogonal to RBME.

That being said, I would still argue that RBME would give statisticians concerned about NHST a useful tool with which they may now more directly effect their sought-after reduction or elimination of its inappropriate use. Authors will now know that publication rests *solely* on sound research methodology, including statistical methodology, because that is *all* the reviewers will see in making their initial decision on publication. Thus, they may

now curiously find themselves being asked by some statisticians reviewing their submitted manuscripts to justify why they propose using NHST, given its widely cited methodological problems, which means the investigators will have to learn about what these problems are as well as alternatives to NHST, and address these issues if they want to get published. Instead of feeling that that they have to get $p < .05$, they will instead find themselves being required to explain why they seek to do so. So to the degree that students; investigators; and journal authors, reviewers, and editors can now be convinced of methodological flaws in NHST, the sooner it should wither and die, to the degree it is indeed invalid and *should* die. Without the forced restriction of an explicit ban on NHST, it will nonetheless become implicitly banned to the extent warranted on the basis of relevant methodological considerations, and *only* those considerations. However, if some residual applications and supplementary value in some contexts can be found for NHST, and effectively argued for in the Methods-Statistics section of a manuscript, then to that degree it might still be retained, and theoretically it should be, to that degree. Although well-intentioned journal bans on p values have had the salutary effect of bringing greater attention to NHST problems, long-term bans would have the downside of prohibiting exploration of any possibly valid uses of NHST.

With RBME, the arguments for and against NHST will now at least have to be debated by everyone on the focused, level playing field of methodology where it should be, without any distracting, irrelevant, and biasing practical concerns regarding the *results* of NHST that confuses the issue. (Note that under RBME, if p values are reported, they will theoretically uniformly span the entire 0–1.00 range.) With RBME, there will no longer be a widespread, perceived *practical* importance and *imperative* attached to the use of NHST, whereby investigators feel they *must* use NHST because they *must* meet an arbitrary $p < .05$ criterion in their results because they mistakenly think *some* criterion like this in the *results* must serve as a minimal requirement for publication, and they must get publications to advance their careers. After decades of seemingly futile protestations by many statisticians about this practice, it will finally come to an end. That's teeth enough for me.

Response to Dr. Kline's critique

Dr. Kline's (2017/in press) criticisms primarily center around two issues:

One is his skepticism that RBME would not do much "about the rampant overreliance on significance testing" (p. 256). My response to this point is essentially the same as I made in addressing Dr. Marks's and Dr. Grice's similar points, and so I do not repeat it here.

The other issue is Dr. Kline's suspicion that at the second stage of RBME, a reviewer's decision regarding publication may still be influenced by how the results turned out, and certainly this could happen, consciously or unconsciously. However, note that 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, self-contradictory* manner, that is, to subsequently want to reject that for which one has already formally stated a preliminary 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 meritorious. Furthermore, by the rules of RBME, a second-stage rejection will require the reviewer to justify to the editor that his or 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.

References

- Grice JW. Comment on Locascio's results blind manuscript evaluation proposal. *Basic and Applied Social Psychology*. 2017; 39(5):254–255.
- Hyman MR. Can “results blind manuscript evaluation” assuage “publication bias”? *Basic and Applied Social Psychology*. 2017; 39(5):247–251.
- Kline RB. Comment on Locascio, results blind science publishing. *Basic and Applied Social Psychology*. 2017; 39(5):256–257.
- Marks M. Commentary on Locascio. *Basic and Applied Social Psychology*. 2017; 39(5):252–253.
- Trafimow D. Editorial. *Basic and Applied Social Psychology*. 2014; 36(1):1–2.
- Trafimow D, Earp BD. Null hypothesis significance testing and Type I error: the domain problem. *New Ideas in Psychology*. 2017; 45:19–27.
- Trafimow D, Marks M. Editorial. *Basic and Applied Social Psychology*. 2015; 37(1):1–2.