

FROM THE EDITORS

DATA OVERLAP POLICIES AT *AMJ*

One question commonly posed at “Meet the Editors” panels concerns the submission of manuscripts whose data overlap with previously published or submitted papers. This question is understandable, given that norms surrounding such issues have evolved over the years, with journals varying in their handling of them. Unlike the procedures governing plagiarism screening (see the August 2012 “From the Editors” [vol. 55: 749–751]), *AMJ*’s data overlap policies are not new. That said, our team has taken steps to bring more clarity and specificity to those policies. The purpose of this “From the Editors” is therefore to reiterate the *Journal*’s expectations for, and handling of, submissions with data overlap.

Before doing so, I should emphasize two points. First, the salience and handling of data overlap issues vary to some degree when submissions are micro vs. macro, quantitative vs. qualitative, lab vs. field, and reliant on primary data vs. secondary data. As a “big tent” journal, *AMJ* must use policies and practices that are able to govern all of the submissions that it receives. Second, a journal’s data overlap policies affect three distinct stakeholders: the action editor who handles a submission, its reviewers, and the eventual reader of the published article. Our policies revolve primarily around the action editor and the eventual reader, though I will return to a discussion of reviewers at the conclusion of this editorial.

SUBMITTING A MANUSCRIPT WITH DATA OVERLAP

AMJ’s policies on these matters become relevant the moment an author submits a manuscript whose data overlap with those of a previously published or submitted paper. It may be that this submission was “part of the plan” from the outset of the author’s research project, or it may be that the desire for the submission emerged once an acceptance or rejection of another manuscript occurred (Kirkman & Chen, 2011). Regardless, submitters are asked two questions when uploading their manuscript into *AMJ*’s system:

Have you previously submitted another manuscript to *AMJ* that uses either some of the same variables from this data collection effort, or that uses some of the same cases/observations from this data collection effort? If so, please describe in detail the precise nature of the data overlap in your cover letter and please attach that previous manuscript in the Cover Letter field.

Have you published (in *AMJ*, in another journal—including “in press” and “conditional accept” stages—or in a book) another publication that uses either some of the same variables from this data collection effort, or that uses some of the same cases/observations from this data collection effort? If so, please describe in detail the precise nature of the data overlap in your cover letter and please attach that previous publication in the Cover Letter field.

These questions define data overlap in the broadest possible terms—overlap of either variables or cases/observations from the same data collection effort. From an author’s perspective, the key is to respond to these questions in as transparent a manner as possible (Kirkman & Chen, 2011). If an author perceives some definitional “gray area” that casts doubt on whether these questions should be answered in the affirmative, the best course of action is to fully explain those issues in the cover letter. This is a critical point because the failure to answer these two questions accurately may result in a “desk rejection,” in which the manuscript is rejected without completion of the double-blind review process. Such inaccuracies may come to light in a number of ways, including the identification of overlapping sample descriptions in web-searchable manuscripts by our CrossCheck screening (see the August 2012 “From the Editors” [vol. 55: 749–751]).

Once the data overlap has been stipulated and explained in a submitting author’s cover letter, the onus falls on the editor to determine whether the submission is “new enough” to be placed under review at the *Journal*. In practice, that determination relies on the same three questions used to govern the submission of previously rejected manuscripts (see the February 2009 “From the Editors”

[vol. 52: 9–10]). First, does the submission address modified or new research questions? Second, does the submission use new theoretical arguments? Third, does the submission use additional or new data to test the proposed relationships? If the answer to all three questions seems to be “yes,” then the manuscript is placed under review, using the typical double-blind review process.

All three questions possess some subjectivity, of course, and the determination of whether a submission is “new enough” is very much case by case. Some of that subjectivity is captured in the Academy’s *Ethics of Research and Publishing Video Series*, in the “Slicing the Data in Publishing” installment (Academy of Management, 2012). I have sometimes relied on associate editors to help me determine whether a submission should be placed under review, particularly with macro or qualitative manuscripts. For hypothetico-deductive papers, the focus is often on the degree of overlap in relevant independent and dependent variables. For inductive papers, relevant concerns include the uniqueness of the data examples and the novelty of the theoretical insights that are derived.

Before leaving this issue, I would like to stress one point: two manuscripts with overlapping data may be different enough from one another without being strong enough to independently survive *AMJ*’s review process. The significance, novelty, and scope of a manuscript’s topic are strong predictors of the manuscript’s ultimate acceptance (see the June 2011 “From the Editors” [vol. 54: 432–435]). With hypothetico-deductive papers, the proper specification of one’s model becomes another strong predictor (see the August 2011 “From the Editors” [vol. 54: 657–660]). One of my most frustrating experiences as an action editor on organizational behavior papers is rejecting a submission because it is “too small” conceptually and methodologically, only to later reject a second submission with overlapping data on the very same grounds. Had the two submissions been combined, or left as one originally, then the resulting manuscript might have been published and made a useful contribution to the literature (see Chen [2011] for a similar observation).

HAVING A MANUSCRIPT WITH DATA OVERLAP ACCEPTED

The discussion above focused primarily on informing the action editor for a submitted manuscript of data overlap. If the submission eventually

gets accepted, then a second stakeholder becomes relevant: the eventual reader of the publication. If the data overlap with a previously published article, then the reader clearly needs to know about that overlap. Such knowledge helps the reader understand the larger research stream, gauge the incremental contribution offered by the new article, and draw his or her own conclusions about any “data slicing” issues. Moreover, if the reader is a meta-analyst, then he or she will realize that effect sizes from the article are not independent from the effect sizes in the overlapping article.

A description of any data overlap gets added to a submission once it has been conditionally accepted. At this stage, the authors typically work on remaining changes solely with their action editor, with the reviewers removed from the process. The current template for conditional acceptance letters includes the following passage:

When you submitted your manuscript, you were asked to indicate to the action editor whether you had previously published another manuscript that uses either some of the same variables from this data collection effort, or that uses some of the same cases/observations from this data collection effort. You were also asked to explain, in detail, the nature of that overlap. If you indicated that there is some overlap in variables or cases/observations, please now detail the precise nature of that overlap for the reader in your Methods section, including all necessary citations and references. Those passages are important to give the reader a sense of where this paper fits into your larger stream of research, and are also critical to future meta-analysts who may need to gauge the nonindependence of findings across studies. Please reproduce the text inserted into your Methods section in your responses document, so that the action editor can inspect it clearly and comprehensively.

This step essentially incorporates many of the details from the submission’s cover letter into the manuscript proper, now that the need to protect author anonymity has been removed. Indeed, such details allow the eventual reader to form his or her own impressions of the “newness” of the published article, relative to the other work drawn from the data collection effort.

CONCLUSIONS

Having reiterated (and more fully described) *AMJ*’s handling of data overlap issues, I should conclude by speaking to two remaining questions.

First, why does our submission process not ask about data overlap with papers currently under review at journals other than *AMJ*? Although authors are, of course, invited to be transparent about such manuscripts in their cover letters, we do not ask about them explicitly because (a) they may be rejected by those outlets, and (b) their constitutive elements may change in a way that alters the overlap. If the review process continues to unfold favorably for both the *AMJ* submission and the other paper, eventually the following will tend to happen: The authors will submit a revised version of their manuscript to *AMJ*, at which point they will be “reasked” the two questions quoted above. If the other paper has been “conditionally accepted” or become “in press” since their last submission, then a previously “no” response to a data overlap question will become a “yes.” At that point, the very same calculation will occur on the part of the editor, with the operative question being whether the *AMJ* resubmission is “new enough” relative to the now published work, in terms of its research question, its theoretical arguments, and its data.

Second, why are *AMJ* reviewers not informed about overlapping data, given that they are one of the relevant stakeholders in the matter? This is an issue that has been debated by our editorial team, and we acknowledge that reasonable editors can disagree about it. Indeed, some top-tier management journals do—understandably—inform reviewers about the exact nature of the data overlap in a manuscript under review by sharing many of the details that would be included in its author’s

cover letter. Ultimately our belief is that any “data overlap discount” should, if relevant, be applied by the action editor for a paper rather than a reviewer. Reviewers may vary in their responses to data overlap, reducing the interrater reliability of their recommendations to the action editor, and may also be differentially affected by a weakening of author anonymity. For their part, action editors possess more complete information about submissions and relevant journal policy, and they will have encountered submissions with data overlap much more frequently. Our hope is that this policy will allow reviewers to do what they do best: Assess, critique, and improve the theoretical, empirical, and practical contributions made by submissions to the *Journal*.

Jason A. Colquitt
University of Georgia

REFERENCES

- Academy of Management. 2012. *Ethics video series: Plagiarism*. Available at: <http://www.youtube.com/watch?v=gGo9oK8v0bc&list=PL65B059BC12E75502&index=3>. Accessed December 16.
- Chen, X.-P. 2011. Author ethical dilemmas in the research publication process. *Management and Organization Review*, 7: 423–432.
- Kirkman, B. L., & Chen, G. 2011. Maximizing your data or data slicing? Recommendations for managing multiple submissions from the same dataset. *Management and Organization Review*, 7: 433–446.