© Academy of Management Journal
2012, Vol. 55, No. 1, 8–12.
http://dx.doi.org/10.5465/amj.2012.4001

FROM THE EDITORS

PUBLISHING IN *AMJ*—PART 5: CRAFTING THE METHODS AND RESULTS

Editor's Note:

This editorial continues a seven-part series, "Publishing in AMJ*," in which the editors give suggestions and advice for improving the quality of submissions to the* Journal*. The series offers "bumper to bumper" coverage, with installments ranging from topic choice to crafting a Discussion section. The series will continue in April with "Part 6: Discussing the Implications." - J.A.C.*

Once the arduous, but exciting, work of selecting an intriguing and appropriate topic, designing and executing a sound data collection, crafting a compelling "hook," and developing a solid theory is finished, it is tempting to sit back, relax, and cruise through the Methods and Results. It seems straightforward, and perhaps a little mundane, to report to the readers (1) how and why the data were obtained; (2) how the data were analyzed and what was found. Indeed, it is unlikely that many readers of *AMJ* have waited with bated breath for an entertaining narrative in this installment of the Publishing in *AMJ* editorial series. If we fall short of being compelling, therefore, we hope to at least be informative.

As authors ourselves, we have, admittedly, succumbed to the temptation of relaxing our concentration when it is time to write these sections. We have heard colleagues say that they pass off these sections to junior members of their research teams to "get their feet wet" in manuscript crafting, as though these sections were of less importance than the opening, hypothesis development, and Discussion sections. Perhaps this is so. But as members of the current editorial team for the past two years, we have come face-to-face with the reality that the Methods and Results sections, if not the most critical, often play a major role in how reviewers evaluate a manuscript. Instead of providing a clear, detailed account of the data collection procedures and findings, these sections often leave reviewers perplexed and raise more questions than they answer about the research procedures and findings that the authors used. In contrast, an effective presentation can have a crucial impact on the extent to which authors can convince their audiences that their theoretical arguments (or parts of them) are supported. High-quality Methods and Results sections also send positive signals about the conscientiousness of the author(s). Knowing that they were careful and rigorous in their preparation of these sections may make a difference for reviewers debating whether to recommend a rejection or a revision request.

To better understand the common concerns raised by reviewers, we evaluated each of our decision letters for rejected manuscripts to this point in our term. We found several issues arose much more frequently in rejected manuscripts than they did in manuscripts for which revisions were requested. The results of our evaluation, if not surprising, revealed a remarkably consistent set of major concerns for both sections, which we summarize as "the three C's": completeness, clarity, and credibility.

THE METHODS

Completeness

In the review of our decision letters, perhaps the most common theme related to Methods sections was that the authors failed to provide a complete description of the ways they obtained the data, the operationalizations of the constructs that they used, and the types of analyses that they conducted. When authors have collected their data—a primary data collection—it is important for them to explain in detail not only what happened, but why they made certain decisions. A good example is found in Bommer, Dierdorff, and Rubin's (2007) study of group-level citizenship behaviors and job performance. We learn in their Methods how the participants were contacted (i.e., on site, by the study's first author), how the data were obtained (i.e., in an on-site training room, from groups of 20–30 employees), what kinds of encouragement for participation were used (i.e., letters from both the company president and the researchers), and who reported the information for different constructs in the model (i.e., employees, supervisors, and managers of the supervisors). In addition, these authors reported other relevant pieces of information about their data collection. For example, they noted that employees and their supervisors were never scheduled to complete their questionnaires in the same room together. In addition, they reported a system of "checks and balances" to make sure supervisors reported performance for all of their direct reports. Providing these details, in addition to a full description of the characteristics of the analysis sample at the individual and team levels, allows reviewers to evaluate the strengths and weaknesses of a research design. Although it is reasonable to highlight the strengths of one's research, reporting sufficient details on the strengths and potential weaknesses of the data collection is preferred over an approach that conceals important details, because certain compromises or flaws can also yield advantages. Consider the example of data collected with a snowball sampling approach in two waves separated by a few months. A disadvantage of this approach would likely be that the sample matched over the two waves will be smaller than the sample resulting if the researchers only contact wave 1 participants to participate in wave 2. But, this approach also has certain advantages. In particular, large numbers of one-wave participants (i.e., those that participated either in the first wave or the second wave) can be used to address response bias and representativeness issues straightforwardly.

In many other cases, the data for a study were obtained from archival sources. Here a researcher may not have access to all the nitty-gritty details of the data collection procedures, but completeness in reporting is no less important. Most, if not all, archival data sets come with technical reports or usage manuals that provide a good deal of detail. Armed with these, the researcher can attempt to replicate the detail of the data collection procedures and measures that is found in primary data collections. For a good example, using the National Longitudinal Survey and Youth Cohort (NLSY79), see Lee, Gerhart, Weller, and Trevor (2008). For other archival data collections, authors construct the dataset themselves, perhaps by coding corporate filings, media accounts, or building variables from other sources. In these cases, a complete description of how they identified the sample, how many observations were lost for different reasons, how they conducted the coding, and what judgment calls were made are necessary.

Regardless of the type of data set a researcher has used, the goals in this section are the same. First, authors should disclose the hows, whats, and whys of the research procedures. Including an Appendix with a full list of measures (and items, where appropriate), for example, is often a nice touch. Second, completeness allows readers to evaluate the advantages and disadvantages of the approach taken, which on balance, creates a more positive impression of the study. Third, a primary goal of the Methods section should be to provide sufficient information that someone could replicate the study and get the same results, if they used exactly the same procedure and data. After reading the Methods section, readers should have confidence that they could replicate the primary data collection or compile the same archival database that the authors are reporting.

Clarity

Far too often, authors fail to clearly explain what they have done. Although there are many potential examples, a typical, very common, problem concerns descriptions of measures. Reviewers are often concerned with language such as "we adapted items" or "we used items from several sources." Indeed, not reporting *how* measures were adapted was the modal issue related to measurement in the evaluation of our decision letters. Ideally, authors can avoid these problems simply by using the full, validated measures of constructs when they are available. When this is not possible, it is imperative to provide a justification for the modifications and, ideally, to provide additional, empirical validation of the altered measures. If this information is not initially included, reviewers will invariably ask for it; providing the information up front improves the chances of a revision request.

Another very common clarity issue concerns the justification for variable coding. Coding decisions are made in nearly every quantitative study, but are perhaps most frequently seen in research involving archival data sets, experimental designs, and assignment of numerical codes based on qualitative responses. For example, Ferrier (2001) used structured content analysis to code news headlines for measures of competitive attacks. In an excellent example of clarity, Ferrier described in an organized fashion and with straightforward language how the research team made the coding decisions for each dimension and how these decisions resulted in operationalizations that matched the constitutive definitions of the competitive attack dimensions.

Credibility

Authors can do several uncomplicated things to enhance perceptions of credibility in their Methods sections. First, it is important to address why a particular sample was chosen. Reviewers often question why a particular sample was used, especially when it is not immediately obvious why the phenomenon of interest is important in the setting used. For example, in Tangirala and Ramanujam's study of voice, personal control, and organizational identification, the authors opened the Methods by describing why they chose to sample front-line hospital nurses to test their hypotheses, noting (1) "they are well positioned to observe early signs of unsafe conditions in patient care and bring them to the attention of the hospital" and (2) "there is a growing recognition that the willingness of nurses to speak up about problems in care delivery is critical for improving patient safety and reducing avoidable medical errors (such as administration of the wrong drug), a leading cause of patient injury and death in the United States" (2008: 1,193). Second, it is always good practice to summarize the conceptual definition of a construct before describing the measure used for it. This not only makes it easier for readers—they don't have to flip back and forth in the paper to find the constitutive definitions—but when done well will lessen reader concerns about whether the theory a paper presents matches the tests that were conducted. Third, it is always important to explain why a particular operationalization was used. For example, organizational performance has numerous dimensions. Some may be relevant to the hypotheses at hand, and others are not. We have often seen authors "surprise" reviewers by introducing certain dimensions with no justification. In cases in which alternative measures are available, authors should report what other measures they considered and why they were not chosen. If alternative measures are available in the data set, it is often a good idea to report the findings obtained when those alternative measures were used. Fourth, it is crucial to justify model specification and data analysis approaches. We have often seen authors include control variables without sufficiently justifying why they should be controlled for. For some types of data, multiple possible methods for analysis exist. Authors need to justify why one method rather than the other(s) was used. Panel data, for example, can be analyzed using fixed-effect models or random-effect models. Multiple event history analysis methods can analyze survival data. Each method has its specific assumption(s). In some cases, additional analysis is warranted to make the choice (for example, doing a Hausman test to choose between fixed- and random-effect models for panel data).

THE RESULTS

Completeness

Effectively writing a Results section is not an easy task, especially when one's theoretical framework and/or research design is complex, making completeness all the more important. For starters, including a table for means, standard deviation, and correlations is a piece of "low-hanging fruit." The information in this table may not have directly tested hypotheses, yet it paints an overall picture of the data, which is critical for judging the credibility of findings. For example, high correlations between variables often raise concerns about multicollinearity. A large standard deviation relative to the mean of a variable can raise concerns about outliers. Indeed it is a good practice to check data ranges and outliers in the process of data analyses so as to avoid having significant findings mainly driven by a few outliers. Distributional properties of variables (such as means and minimum and maximum values) reported in a table are informative by themselves. For example, in a study on CEO succession, means of variables that measured different types of CEO successions can tell the distribution of new CEOs in the sample recruited from different sources. These distributional properties describe the phenomenon of CEO successions and have important practical implications.

In reporting results, it is important to specify the unit of analysis, sample size, and dependent variable used in each model. This is especially crucial when such information varies across models. Take Arthaud-Day, Certo, Dalton, and Dalton (2006) as an example. These authors examined executive and director turnover following corporate financial restatements. They had four dependent variables: CEO turnover, CFO turnover, outside director turnover, and auditing commitment member turnover. In models of CEO and CFO turnover, because they were able to identify the month of the turnover, they constructed the data using "CEO/CFO" as the unit of analysis and used a Cox model to examine the timing of the executive turnover. The sample size of the model on CEO turnover was 485, and the sample size of the model on CFO turnover was 407. In comparison, in examining turnover of outside directors and audit committee members, because Arthaud-Day and her colleagues were unable to determine the month in which outside directors and audit committee members left office, they constructed the data using director/auditing committee member-year as the unit of analysis and used logistic regression to examine the likelihood of their turnover. The sample size of the model on outside director turnover was 2,668, and the sample size for auditing committee member turnover was 1,327. The take-away here is that careful descriptions such as those Arthaud-Day and colleagues provided help readers calibrate their interpretations of results and prevent reviewers from raising questions about clarification.

Clarity

The purpose of a Results section is to answer the research questions that have been posed and provide empirical evidence for the hypotheses (or note that evidence is lacking). We often see, however, that authors do not relate their findings to the study's hypotheses. We also see that authors report the results in the Results section, but discuss their linkage with hypotheses in the Discussion section or, conversely, begin to discuss the implications of the findings in the Results prematurely, rather than doing this in the Discussion. In these cases, the authors fail to describe what the results indicate with respect to the focal topic of the study in a clear manner. To avoid this problem, it helps to summarize each hypothesis before reporting the related results. Try this format: "Hypothesis X suggests that . . . We find that . . . in model . . . in Table . . . Thus, Hypothesis X is (or isn't) supported." Although this format may sound mechanical or even boring, it is a very effective way to clearly report results (see also Bem, 1987). We encourage and welcome authors to experiment with novel and *clear* ways to present results. We also suggest that authors report the results associated with their hypotheses in order, beginning with the first hypothesis and continuing sequentially to the last one, unless some compelling reasons suggest that a different order is better.

In many studies, the results do not support all the hypotheses. Yet results that are not statistically significant and those with signs opposite to prediction are just as important as those that are supported. However, as one editor noted, "If the results are contrary to expectations, I find authors will often try to 'sweep them under the rug.'" Of course, reviewers will catch this immediately. Needless to say, sometimes such results reflect inadequate theorizing (e.g., the hypotheses are wrong, or at least there are alternative arguments and predictions). Other times, however, unsupported results are great fodder for new, critical thinking in a Discussion section. The point is that all results—significant or not, supporting or opposite to hypotheses—need to be addressed directly and clearly.

It is also a good practice to reference variables across sections in the same order—for example, describe their measures in the Methods section, list them in tables, and discuss results in the Results section all in the same order. Such consistency improves the clarity of exposition and helps readers to both follow the manuscript and find information easily. It also provides authors with a checklist so that they will remember to include relevant information (e.g., a variable included in the models is not mentioned in the Methods section and/or in the correlation matrix).

Credibility

Although every part of a paper plays an important role in helping or hurting its credibility (e.g., adequate theorizing and rigorous research design), there are some things authors can do in their Results sections to enhance the perceived credibility of findings. First, it is crucial to demonstrate to readers why one's interpretations of results are correct. For example, a negative coefficient for an interaction term may suggest that the positive effect of the predictor became weaker, or disappeared, or even became negative as the value of the moderator increased. Plotting a significant interaction effect helps one visualize the finding and thus demonstrate whether the finding is consistent with the intended hypothesis. Aiken and West (1991) provided some "golden rules" on how to plot interaction effects in regressions. Beyond these, determining whether the simple slopes are statistically significant is often important in assessing whether one's results fully support hypotheses; techniques developed by Preacher, Curran, and Bauer (2006) are helpful in these calculations.

Second, if alternative measurements, methods, and/or model specifications could be used for a study, but authors only report results using one possible choice, readers may have the impression that the authors "cherry-picked" findings that were consistent with the hypotheses. Supplementary analyses and robustness checks can address these concerns. For example, Tsai and Ghoshal (1998) examined the value creation role of a business unit's position in intrafirm networks. Although they proposed the hypotheses at the individual business unit level, they generated several measures of business units' attributes from data at the dyadic level. These steps raised some concerns about level of analysis and the reliability of the results. To address these concerns, they also analyzed data at the dyadic level and obtained consistent results.

Third, even if a result is statistically significant, readers may still ask, So what? A statistically significant effect is not necessarily a practically important effect. Authors typically discuss the practical implications of a study in their Discussion; they can, however, conduct and report additional analyses in Results to demonstrate the practical relevance of findings. A good example is found in Barnett and King's (2008) study of spillover harm. These authors stated the following Hypothesis 1: "An error at one firm harms other firms in the same industry" (Barnett & King, 2008: 1,153). In addition to reporting the statistical significance of the predictor, the authors provided information to communicate the average scale of such spillovers. They reported that "following an accident that injured an average number of employees (3.5), a chemical firm with operations in the same industry as that in which an accident occurred could expect to lose 0.15 percent of its stock price" and that "after an accident that caused the death of an employee, the firm could expect to lose an additional 0.83 percent" (Barnett & King, 2008: 1,160). In other cases, authors may want to discuss the implications of small effect sizes, perhaps by noting how difficult it is to explain variance in a given dependent variable or, in the case, of an experiment, noting that a significant effect was found even though the manipulation of the independent variable was quite minimal (Prentice & Miller, 1992).

Conclusions

Crafting Methods and Results sections may not sound exciting or challenging. As a result, authors tend to pay less attention in writing them. Sometimes these sections are delegated to the junior members of research teams. However, in our experience as editors, we find that these sections often play a major, if not a critical, role in reviewers' evaluations of a manuscript. We urge authors to take greater care in crafting these sections. The three-C rule—completeness, clarity, and credibility—is one recipe to follow in that regard.

Yan (Anthea) Zhang
Rice University

Jason D. Shaw
University of Minnesota

REFERENCES

Aiken, L. S., & West, S. G. 1991. ***Multiple regression: Testing and interpreting interactions.*** Newbury Park, CA: Sage.

Arthaud-Day, M. L., Certo, S. T., Dalton, C. M., & Dalton, D. R. 2006. A changing of the guard: Executive and director turnover following corporate financial restatements. ***Academy of Management Journal,*** 49: 1119–1136.

Barnett, M. L., & King, A. A. 2008. Good fences make good neighbors: A longitudinal analysis of an industry self-regulatory institution. ***Academy of Management Journal,*** 51: 1150–1170.

Bem, D. J. 1987. Writing the empirical journal article. In M. P. Zanna & J. M. Darley, (Eds.), ***The compleat academic: A practical guide for the beginning social scientist:*** 171–201. New York: Random House.

Bommer, W. H., Dierdorff, E. C., & Rubin, R. S. 2007. Does prevalence mitigate relevance? The moderating effect of group-level OCB on employee performance. ***Academy of Management Journal,*** 50: 1481–1494.

Ferrier, W. J. 2001. Navigating the competitive landscape: The drivers and consequences of competitive aggressiveness. ***Academy of Management Journal,*** 44: 858–877.

Lee, T. H., Gerhart, B., Weller, I., & Trevor, C. O. 2008. Understanding voluntary turnover: Path-specific job satisfaction effects and the importance of unsolicited job offers. ***Academy of Management Journal,*** 51: 651–671.

Preacher, K. J., Curran, P. J., & Bauer, D. J. 2006. Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. ***Journal of Educational and Behavioral Statistics,*** 31: 437–448.

Prentice, D. A., & Miller, D. T. 1992. When small effects are impressive. ***Psychological Bulletin,*** 112: 160–164.

Tangirila, S., & Ramanujam, R. 2008. Exploring nonlinearity in employee voice: The effects of personal control and organizational identification. ***Academy of Management Journal,*** 51: 1189–1203.

Tsai, W., & Ghoshal, S. 1998. Social capital and value creation: The role of intrafirm networks. ***Academy of Management Journal,*** 41: 464–474.