

FROM THE EDITORS

PUBLISHING IN *AMJ*—PART 2: RESEARCH DESIGN

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 October with "Part 3: Setting the Hook."- J.A.C.

Most scholars, as part of their doctoral education, take a research methodology course in which they learn the basics of good research design, including that design should be driven by the questions being asked and that threats to validity should be avoided. For this reason, there is little novelty in our discussion of research design. Rather, we focus on common design issues that lead to rejected manuscripts at *AMJ*. The practical problem confronting researchers as they design studies is that (a) there are no hard and fast rules to apply; matching research design to research questions is as much art as science; and (b) external factors sometimes constrain researchers' ability to carry out optimal designs (McGrath, 1981).

Access to organizations, the people in them, and rich data about them present a significant challenge for management scholars, but if such constraints become the central driver of design decisions, the outcome is a manuscript with many plausible alternative explanations for the results, which leads ultimately to rejection and the waste of considerable time, effort, and money. Choosing the appropriate design is critical to the success of a manuscript at *AMJ*, in part because the fundamental design of a study cannot be altered during the revision process. Decisions made during the research design process ultimately impact the degree of confidence readers can place in the conclusions drawn from a study, the degree to which the results provide a strong test of the researcher's arguments, and the degree to which alternative explanations can be discounted. In reviewing articles that have been rejected by *AMJ* during the past year, we identified three broad design problems that were common sources of rejection: (a) mismatch between research question and design, (b) measurement and operational issues (i.e., construct validity), and (c) inappropriate or incomplete model specification.

Matching Research Question and Design

Cross-sectional data. Use of cross-sectional data is a common cause of rejection at *AMJ*, of both micro

and macro research. Rejection does not happen because such data are inherently flawed or because reviewers or editors are biased against such data. It happens because many (perhaps most) research questions in management implicitly—even if not framed as such—address issues of change. The problem with cross-sectional data is that they are mismatched with research questions that implicitly or explicitly deal with causality or change, strong tests of which require either measurement of some variable more than once, or manipulation of one variable that is subsequently linked to another. For example, research addressing such topics as the effects of changes in organizational leadership on a firm's investment patterns, the effects of CEO or TMT stock options on a firm's actions, or the effects of changes in industry structure on behavior implicitly addresses causality and change. Similarly, when researchers posit that managerial behavior affects employee motivation, that HR practices reduce turnover, or that gender stereotypes constrain the advancement of women managers, they are also implicitly testing change and thus cannot conduct adequate tests with cross-sectional data, regardless of whether that data was drawn from a pre-existing data base or collected via an employee survey. Researchers simply cannot develop strong causal attributions with cross-sectional data, nor can they establish change, regardless of which analytical tools they use. Instead, longitudinal, panel, or experimental data are needed to make inferences about change or to establish strong causal inferences. For example, Nyberg, Fulmer, Gerhart, and Carpenter (2010) created a panel set of data and used fixed-effects regression to model the degree to which CEO-shareholder financial alignment influences future shareholder returns. This data structure allowed the researchers to control for cross-firm heterogeneity and appropriately model how changes in alignment within firms influenced shareholder returns.

Our point is not to denigrate the potential usefulness of cross-sectional data. Rather, we point out the importance of carefully matching research design to research question, so that a study or set of studies

is capable of testing the question of interest. Researchers should ask themselves during the design stage whether their underlying question can actually be answered with their chosen design. If the question involves change or causal associations between variables (any mediation study implies causal associations), cross-sectional data are a poor choice.

Inappropriate samples and procedures. Much organizational research, including that published in *AMJ*, uses convenience samples, simulated business situations, or artificial tasks. From a design standpoint, the issue is whether the sample and procedures are appropriate for the research question. Asking students with limited work experience to participate in experimental research in which they make executive selection decisions may not be an appropriate way to test the effects of gender stereotypes on reactions to male and female managers. But asking these same students to participate in a scenario-based experiment in which they select the manager they would prefer to work for may present a good fit between sample and research question. Illustrating this notion of matching research question with sample is a study on the valuation of equity-based pay in which Devers, Wiseman, and Holmes (2007) used a sample of executive MBA students, nearly all of whom had experience with contingent pay. The same care used in choosing a sample needs to be taken in matching procedures to research question. If a study involves an unfolding scenario wherein a subject makes a series of decisions over time, responding to feedback about these decisions, researchers will be well served by collecting data over time, rather than having a series of decision and feedback points contained in a single 45 minute laboratory session.

Our point is not to suggest that certain samples (e.g., executives or students) or procedures are inherently better than others. Indeed, at *AMJ* we explicitly encourage experimental research because it is an excellent way to address questions of causality, and we recognize that important questions—especially those that deal with psychological process—can often be answered equally well with university students or organizational employees (see *AMJ*'s August 2008 *From the Editors* [vol. 51: 616–620]). What we ask of authors—whether their research occurs in the lab or the field—is that they match their sample and procedures to their research question and clearly make the case in their manuscript for why these sample or procedures are appropriate.

Measurement and Operationalization

Researchers often think of validity once they begin operationalizing constructs, but this may be too late. Prior to making operational decisions, an au-

thor developing a new construct must clearly articulate the definition and boundaries of the new construct, map its association with existing constructs, and avoid assumptions that scales with the same name reflect the same construct and that scales with different names reflect different constructs (i.e., jingle jangle fallacies [Block, 1995]). Failure to define the core construct often leads to inconsistency in a manuscript. For example, in writing a paper, authors may initially focus on one construct, such as organizational legitimacy, but later couch the discussion in terms of a different but related construct, such as reputation or status. In such cases, reviewers are left without a clear understanding of the intended construct or its theoretical meaning. Although developing theory is not a specific component of research design, readers and reviewers of a manuscript should be able to clearly understand the conceptual meaning of a construct and see evidence that it has been appropriately measured.

Inappropriate adaptation of existing measures. A key challenge for researchers who collect field data is getting organizations and managers to comply, and survey length is frequently a point of concern. An easy way to reduce survey length is to eliminate items. Problems arise, however, when researchers pick and choose items from existing scales (or rewrite them to better reflect their unique context) without providing supporting validity evidence. There are several ways to address this problem. First, if a manuscript includes new (or substantially altered measures), all the items should be included in the manuscript, typically in an appendix. This allows reviewers to examine the face validity of the new measures. Second, authors might include both measures (the original and the shortened versions) in a subsample or in an entirely different sample as a way of demonstrating high convergent validity between them. Even better would be including several other key variables in the nomological network, to demonstrate that the new or altered measure is related to other similar and dissimilar constructs.

Inappropriate application of existing measures. Another way to raise red flags with reviewers is to use existing measures to assess completely different constructs. We see this problem occurring particularly among users of large databases. For example, if prior studies have used an action such as change in format (e.g., by a restaurant) as a measure of strategic change, and a submitted paper uses this same action (change in format) as a measure of organizational search, we are left with little confidence that the authors have measured their intended construct. Given the cumulative and incremental nature of the research process, it is critical

that authors establish both the uniqueness of their new construct, how it relates to existing constructs, and the validity of their operationalization.

Common method variance. We see many rejected *AMJ* manuscripts in which data are not only cross-sectional, but are also assessed via a common method (e.g., a survey will have multiple predictor and criterion variables completed by a single individual). Common method variance presents a serious threat to interpretation of observed correlations, because such correlations may be the result of systematic error variance due to measurement methods, including rater effects, item effects, or context effects. Podsakoff, MacKenzie, Lee, and Podsakoff (2003) discussed common method variance in detail and also suggested ways to reduce its biasing effects (see also Conway & Lance, 2010).

Problems of measurement and operationalization of key variables in *AMJ* manuscripts have implications well beyond psychometrics. At a conceptual level, sloppy and imprecise definition and operationalization of key variables threaten the inferences that can be drawn from the research. If the nature and measurement of underlying constructs are not well established, a reader is left with little confidence that the authors have actually tested the model they propose, and reasonable reviewers can find multiple plausible interpretations for the results. As a practical matter, imprecise operational and conceptual definitions also make it difficult to quantitatively aggregate research findings across studies (i.e., to do meta-analysis).

Model Specification

One of the challenges of specifying a theoretical model is that it is practically not feasible to include every possible control variable and mediating process, because the relevant variables may not exist in the database being used, or because organizations constrain the length of surveys. Yet careful attention to the inclusion of key controls and mediating processes during the design stage can provide substantial payback during the review process.

Proper inclusion of control variables. The inclusion of appropriate controls allows researchers to draw more definitive conclusions from their studies. Research can err on the side of too few or too many controls. Control variables should meet three conditions for inclusion in a study (Becker, 2005; James, 1980). First, there is a strong expectation that the variable be correlated with the dependent variable owing to a clear theoretical tie or prior empirical research. Second, there is a strong expectation that the control variable be correlated with the hypothesized independent variable(s).

Third, there is a logical reason that the control variable is not a more central variable in the study, either a hypothesized one or a mediator. If a variable meeting these three conditions is excluded from the study, the results may suffer from omitted variable bias. However, if control variables are included that don't meet these three tests, they may hamper the study by unnecessarily soaking up degrees of freedom or bias the findings related to the hypothesized variables (increasing either type I or type II error) (Becker, 2005). Thus, researchers should think carefully about the controls they include—being sure to include proper controls but excluding superfluous ones.

Operationalizing mediators. A unique characteristic of articles in *AMJ* is that they are expected to test, build, or extend theory, which often takes the form of explaining *why* a set of variables are related. But theory alone isn't enough; it is also important that mediating processes be tested empirically. The question of when mediators should be included in a model (and which mediators) needs to be addressed in the design stage. When an area of inquiry is new, the focus may be on establishing a causal link between two variables. But, once an association has been established, it becomes critical for researchers to clearly describe and measure the process by which variable A affects variable B. As an area of inquiry becomes more mature, multiple mediators may need to be included. For example, one strength of the transformational leadership literature is that many mediating processes have been studied (e.g., LMX [Kark, Shamir, & Chen, 2003; Pillai, Schriesheim, & Williams, 1999; Wang, Law, Hackett, Wang, & Chen, 2005]), but a weakness of this literature is that most of these mediators, even when they are conceptually related to each other, are studied in isolation. Typically, each is treated as if it is the unique process by which managerial actions influence employee attitudes and behavior, and other known mediators are not considered. Failing to assess known, and conceptually related mediators, makes it difficult for authors to convince reviewers that their contribution is a novel one.

Conclusion

Although research methodologies evolve over time, there has been little change in the fundamental principles of good research design: match your design to your question, match construct definition with operationalization, carefully specify your model, use measures with established construct validity or provide such evidence, choose samples and procedures that are appropriate to your unique

research question. The core problem with *AMJ* submissions rejected for design problems is not that they were well-designed studies that ran into problems during execution (though this undoubtedly happens); it is that the researchers made too many compromises at the design stage. Whether a researcher depends on existing databases, actively collects data in organizations, or conducts experimental research, compromises are a reality of the research process. The challenge is to not compromise too much (Kulka, 1981).

A pragmatic approach to research design starts with the assumption that most single-study designs are flawed in some way (with respect to validity). The best approach, then, to a strong research design may not lie in eliminating threats to validity (though they can certainly be reduced during the design process), but rather in conducting a series of studies. Each study in a series will have its own flaws, but together the studies may allow for stronger inferences and more generalizable results than would any single study on its own. In our view, multiple study and multiple sample designs are vastly underutilized in the organizational sciences and in *AMJ* submissions. We encourage researchers to consider the use of multiple studies or samples, each addressing flaws in the other. This can be done by combining field studies with laboratory experiments (e.g., Grant & Berry, 2011), or by testing multiple industry data sets to assess the robustness of findings (e.g., Beck, Bruderl, & Woywode, 2008). As noted in *AMJ*'s "Information for Contributors," it is acceptable for multiple study manuscripts to exceed the 40-page guideline.

A large percentage of manuscripts submitted to *AMJ* that are either never sent out for review or that fare poorly in the review process (i.e., all three reviewers recommend rejection) have flawed designs, but manuscripts published in *AMJ* are not perfect. They sometimes have designs that cannot fully answer their underlying questions, sometimes use poorly validated measures, and sometimes have misspecified models. Addressing all possible threats to validity in each and every study would be impossibly complicated, and empirical research might never get conducted (Kulka, 1981). But honestly assessing threats to validity during the design stage of a research effort and taking steps to minimize them—either via improving a single study or conducting multiple studies—will substantially improve the potential for an ultimately positive outcome.

Joyce E. Bono

University of Florida

Gerry McNamara

Michigan State University

REFERENCES

- Beck, N., Bruderl, J., & Woywode, M. 2008. Momentum or deceleration? Theoretical and methodological reflections on the analysis of organizational change. *Academy of Management Journal*, 51: 413–435.
- Becker, T. E. 2005. Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods*, 8: 274–289.
- Block, J. 1995. A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117: 187–215.
- Conway, J. M., & Lance, C. E. 2010. What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, 25: 325–334.
- Devers, C. E., Wiseman, R. M., & Holmes, R. M. 2007. The effects of endowment and loss aversion in managerial stock option valuation. *Academy of Management Journal*, 50: 191–208.
- Grant, A. M., & Berry, J. W. 2011. The necessity of others is the mother of invention: Intrinsic and prosocial motivations, perspective taking, and creativity. *Academy of Management Journal*, 54: 73–96.
- James, L. R. 1980. The unmeasured variables problem in path analysis. *Journal of Applied Psychology*, 65: 415–421.
- Kark, R., Shamir, B., & Chen, G. 2003. The two faces of transformational leadership: Empowerment and dependency. *Journal of Applied Psychology*, 88: 246–255.
- Kulka, R. A. 1981. Idiosyncrasy and circumstance. *American Behavioral Scientist*, 25: 153–178.
- McGrath, J. E. 1981. Introduction. *American Behavioral Scientist*, 25: 127–130.
- Nyberg, A. J., Fulmer, I. S., Gerhart, B., & Carpenter, M. A. 2010. Agency theory revisited: CEO return and shareholder interest alignment. *Academy of Management Journal*, 53: 1029–1049.
- Pillai, R., Schriesheim, C. A., & Williams, E. S. 1999. Fairness perceptions and trust as mediators for transformational and transactional leadership: A two-sample study. *Journal of Management*, 25: 897–933.
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 25: 879–903.
- Wang, H., Law, K. S., Hackett, R. D., Wang, D., & Chen, Z. X. 2005. Leader-member exchange as a mediator of the relationship between transformational leadership and followers' performance and organizational citizenship behavior. *Academy of Management Journal*, 48: 420–432.