

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/261943014>

# Sophistication in Research in Marketing

Article in *Journal of Marketing* · July 2011

DOI:10.2307/41228617

---

CITATIONS

135

---

READS

1,514

3 authors, including:



**Leigh McAlister**

University of Texas at Austin

70 PUBLICATIONS 7,425 CITATIONS

[SEE PROFILE](#)



**Richard Staelin**

Duke University

154 PUBLICATIONS 21,515 CITATIONS

[SEE PROFILE](#)

# Sophistication in Research in Marketing

Over the years, the level of analytical rigor has risen in articles published in marketing academic journals. While, *ceteris paribus*, rigor is desirable, there is a growing sense that rigor has become a, if not the, goal for research in marketing. Consequently, other desirable characteristics, such as relevance, communicability, and simplicity, have been downplayed, to the detriment of the field of marketing. The authors explore this imbalance, setting forth the consequences of overemphasis on rigor for (1) the manuscript review process, (2) PhD programs, (3) hiring, and (4) the tenure and promotion review process. Two surveys of “successful” authors provide empirical support for the conjectures put forth. The authors then identify the causes for this trend and propose some directions to reestablish a better balance between rigor and relevance.

*Keywords:* priorities, communication, relevance, rigor

**M**arketing, almost by definition, is an applied field of inquiry. Moreover, it is boundary spanning in that it interfaces with almost every aspect of the firm. Consequently, research in marketing should be broad in nature and strive to yield insights that potentially help us better understand the consequences of marketing actions and improve the practice of marketing. Of course, it is also important to ensure that the findings of any research are not incorrect because of faulty research procedures. The tension between making sure our field provides insights that are relevant and accessible and the quest to make sure that the insights are not incorrect as a result of faulty analysis has been present for at least 50 years. However, recently there has been a noticeably increased emphasis on the use of (often unnecessarily) complex analyses and an accompanying decrease in emphasis on the importance of the topics explored and the substantive insights that come from these explorations. This analytical complexity has one clear downside: It hinders communication, making any insights associated with the research less accessible to managers as well as to most other academics. Another consequence of increased complexity is that marketing scholars are partitioning themselves into small groups, in which each group is defined by the particular kind of complexity that group favors, leading to fragmentation of the field. We attribute this fragmentation partly to the fact that only a small subset of scholars feels competent to evaluate (or understand) a particular researcher’s theory, methods, and findings and partly because scholars have concluded that being more

technically sophisticated increases their probability of being “successful” academics. This growing complexity of research and fragmentation of the field increases the likelihood that any given piece of academic research in marketing speaks only to a limited set of marketing academics. It also implies that the research is less likely to examine issues that span boundaries. As a result, academic research in marketing has less impact on both the practice of marketing and overall knowledge development. Importantly, as the field of marketing focuses more attention on rigor, other disciplines are addressing substantively important topics that are inherently concerned with either marketing functions (e.g., supply chains, service quality, pricing, new types of communications) or broad topics related to complex interactions within the firm and between the firm and customers and competitors.<sup>1</sup> The net effect of all this is that our field is becoming increasingly marginalized.

This article considers some of the forces driving the field of marketing toward complexity at the expense of relevance. The tendency to make rigor a goal unto itself during the review process has implications for the types of papers that are published, the type of education provided to PhD students, and the type of marketing faculty being hired and promoted by major business schools. The changes in PhD programs and marketing faculties, in turn, increase the level of complexity expected in academic research in marketing, resulting in self-reinforcing feedback loops. We examine the implications of this complex feedback loop for different stakeholders within the marketing field and provide suggestions of ways to rebalance the trade-off between needed rigor and the quest for relevance.

<sup>1</sup>For example, *Management Science* recently had a call for papers that examine marketing’s various interfaces with other disciplines. More than 150 papers were submitted, a figure that exceeded the annual submission to the department of Marketing within *Management Science*. Moreover, the department exhibited no decrease in normal submissions, implying almost no cannibalization of normal marketing topics.

---

Donald R. Lehmann is George E. Warren Professor of Business, Columbia Business School, Columbia University (e-mail: drl2@columbia.edu). Leigh McAlister is Ed and Molly Smith Chair in Business Administration, McCombs School of Business, University of Texas at Austin (e-mail: leigh.mcalister@mcombs.utexas.edu). Richard Staelin is Edward and Rose Donnell Professor of Business Administration, Fuqua School of Business, Duke University (e-mail: rstaelin@mail.duke.edu). The authors thank Gary Lilien and Bart Weitz for their helpful comments on this article and the 286 respondents to the survey for their cooperation and insights.



## Journal Proliferation

The migration of scholars from the basic disciplines and the resultant incorporation of these disciplines' ideas and concepts in marketing is not the only major change in our field that has led to more emphasis on execution and fragmentation of ideas and problems addressed. From a single journal in 1925 (*Journal of Retailing*), the academic marketing outlets for new ideas expanded fairly modestly to four by 1964 (adding *Journal of Marketing*, *Journal of Advertising Research*, and *Journal of Marketing Research*) and six by 1974 (adding *Journal of Advertising* and *Journal of Consumer Research*). Since then, the number of journals has exploded (for a more complete listing, see Lehmann 2005) and continues to expand. This has encouraged specialization of topics, methods, and language, in effect both recognizing and encouraging a splintering of marketing academics. The net result is that authors are talking to a smaller and smaller subset of scholars who have a common interest in a relatively small set of well-defined problems as well as a common research paradigm and language. The increasing rigidity (if not isolation) of these ever-narrowing domains contributes to the field's move away from big, boundary spanning ideas and toward specialized methodology.

## The Review Process

All these changes have greatly affected the manuscript review process. We discuss the review process in some detail because it plays a pivotal role in what is ultimately published and thus ultimately affects what the field views as "successful" academics and research. In addressing this topic, we acknowledge that we are not the first to discuss these issues, nor is this discussion unique to marketing. For example, discussions about the review process have also appeared in the economics (Ellison 2002) and accounting (Swanson 2004) literatures.

We begin with a discussion of Ellison's (2002) model, in which he posits that reviewers continually try to deduce the current social norm for evaluating overall quality. He argues that the evaluation process naturally shifts toward an overweighting of the execution quality relative to the idea quality (importance and interest of its main ideas). He attributes this to the idea that reviewers overrate the execution quality of their own research and, therefore, based on the reviews they receive of their own work, come to believe that the importance (as viewed by the field) of execution quality is greater than they initially thought. The cycle repeats itself, with the end result being that what Ellison calls "r" quality (executional rigor) dominating "q" quality (relevance and interest). Ellison goes on to argue that this tendency to overemphasize executional rigor also results from reviewers trying to impress the editor by requiring complex revisions and sometimes even imposing executional quality standards above the norm to hold others back.

These changes in the review process have also been discussed in the marketing literature. McAlister (2005) draws attention to the relevance of Ellison's (2002) model for the evolution of the field of marketing. Yadav (2010) documents the shift of *Journal of Marketing* away from concep-

tual articles, and Reibstein, Day, and Wind (2009, pp. 1–2) lay out the implications of marketing's shift in reviewing criteria from idea quality to methodology:

The prevailing research paradigm in most parts of marketing academia is to begin with a new methodology, data set, or a behavioral hypothesis and only then occasionally ask where it might be applied. This reduces the odds of addressing a pressing marketing issue. The resultant conclusions are of some relevance to other researchers, but they offer little guidance for marketing decision making.

Reibstein, Day, and Wind call for a change in the academic marketing research process. In particular, they suggest "that a better approach is to begin with an important problem and bring to it the best combination of methodology, data and theory" (pp. 1–2). This largely reprises the thesis of Lodish (1986), who asks the rhetorical question whether it is better to be vaguely right or precisely wrong.

The shift in reviewing emphasis has had a particularly negative impact on two types of papers. The first are papers that break genuinely new conceptual ground. Such papers are likely to be criticized for their lack of theory. This criticism takes the somewhat absurd position that the new theory must be justified by old theory. Reviewers (and editors) often do not seem to recognize that if theory is new, it often cannot be justified by old theory; rather, new theory may well contradict old theory. (How could Einstein have justified  $E = mc^2$  based on Newton's  $F = ma$ ?) Instead, the operational definition of good execution in terms of theory is a paper that includes a long section describing what others have already published, some of which is typically (1) already widely accepted and (2) obvious. Indeed, a better measure of the potential of a paper to make major new contributions to theory may be the *lack* of references to prior work—in other words, the work stands on its own.

The second type of paper that has been adversely affected is one that presents purely empirical (but interesting) results—that is, new and interesting stylized facts—but is largely devoid of theory. In such cases, the reason often given for rejection is that the paper is not "sophisticated" because it does not test a preexisting theory. Yet Brahe's careful measurements (empirical facts) were the basis for his student Kepler's laws (a new theory). Likewise, the initial work on the mere measurement effect (Fitzsimons and Morwitz 1996) was basically a set of new stylized facts that spawned a stream of research aimed at better understanding why this effect occurred. Another example of a new and interesting stylized fact is the work of Blattberg and Wisniewski (1989), who point out the empirical regularity that promotion of top-tier brands strongly affects lower-tier brands, but in general the converse effect is not observed. This observation has led to others trying to understand why this occurs. Finally, we note in passing that much has been said about whether theory or data come first (i.e., is it ETET or TETE?); see, for example, the discussions of Bass (1995) and Ehrenberg (1995). A more pragmatic and useful view is that it does not matter which comes first, only that each type contributes to future work. This strongly suggests that both types of papers should be valued by our profession and are equally important to the advancement of knowledge. A

corollary to this is that it is not clear that requiring all papers to be strong on theory and empirical results (or on analytical rigor and substantive content) is the best way to advance the field.

A common example of reviewers' overemphasis on the execution quality of a paper is a formulaic adherence to what constitutes "good" research. Take, for example, the "bias" against construct measurement based on only one or two items. Clearly, measurement based on only a few fallible measures has limitations, but it also has benefits. For example, respondent burden is reduced. Furthermore, there is diminishing value in reducing random error through averaging multiple items (which often are merely slight restatements of each other). Formulaic rules such as "you need a minimum of three items per construct" ignore reality: Some constructs can be adequately captured with just one question (e.g., attitude toward ads and brands; Bergkuist and Rossiter 2007), while attitudes toward tenure, war, and so on may be nuanced and multifaceted and may require far more than three items, or even dimensions, to capture them. Other examples are the "rules" that one must (1) use structural equation modeling versus regression when analyzing survey data, (2) run at least three experiments, (3) have interaction effects in addition to main effects, and (4) always have longitudinal data if you want to overcome common method bias. These rules confuse desirability with hard and fast requirements.

## How Successful Authors View the Review Process

We wondered if there is a consensus supporting our conjecture that the review process is slanting the field's research away from big ideas and toward execution. (One hypothesis is that our views are in the minority and that most marketing academics feel the field is moving in the right direction.) Therefore, we broadened our inquiry by sampling "successful" authors, specifically those researchers who had published in *Journal of Marketing (JM)*, *Journal of Marketing Research (JMR)*, *Journal of Consumer Research (JCR)*, or *Marketing Science* between January 2005 and January 2009. (For better or worse, many in the field refer to these as the top four "A" journals in marketing.) We selected this group of authors because we believed that they would be

more likely to have a positive view of the review process, given that they received at least one favorable review outcome. Moreover, they are also likely to be reviewers for those journals and thus part of the "system." We sent e-mails to every author who met this criterion for whom we could obtain an e-mail address. This resulted in a list of 510 authors, of whom 286 completed our survey, for a response rate of 56%. The respondents represented a broad cross-section of the field, with 22% being in the field for 5 or fewer years, 23% for 6–10 years, 26% for 11–20 years, and 24% for 21 or more years (6% of respondents did not respond to this question).

Using the critical incident method, we asked respondents to evaluate two recent reviews, one that they would classify as "useful" and another they would classify as "not useful." Respondents then reported on the nature of requests made for both types of reviews. As we report in Table 1, three kinds of requests were more likely to be associated with "useful" (vs. "not useful") reviews: enhance managerial relevance, provide a stronger theoretical justification, and defeat (rule out) an alternative explanation. Three kinds of requests were more likely to be associated with "not useful" (vs. "useful") reviews: use a different model, use a more complex method, and alter the thrust of the paper. The other two types of requests often mentioned were to extend the exploration to other areas (19%) and collect new data (46%). These were equally likely to be requested in helpful and unhelpful reviews. Thus, it appears that authors perceive reviews as "useful" if the requests focus attention more on the quality of the research question versus the methodology used in the paper. In contrast, authors are more likely to perceive reviews as "not useful" when the review team is concerned with the execution of the research and/or tries to "rewrite" the paper by altering the thrust of the author's intended exploration. The impact of these different requests is also of interest. When those researchers who were asked to use a more complex method or a different model were asked whether their conclusions were altered by the increased complexity, more than 80% indicated that the conclusions did not change; only 19% of the time did these authors report that the more complex analysis gave the reader (or them) new insights. In contrast, they reported the paper became more difficult to read as a consequence of this new analysis 38% of the time. In addi-

**TABLE 1**  
Characteristics of Useful and Not Useful Reviews

	% Indicating That They Received This Request in the "Useful" Review = USEFUL	% Indicating That They Received This Request in the "Not Useful" Review = NOT USEFUL	Index = USEFUL/ NOT USEFUL
<b>Reviewer Request Associated with "Useful" Review</b>			
Enhance managerial relevance	33%	14%	2.36
Provide stronger theoretical justification	60	39	1.54
Defeat an alternative explanation	48	36	1.33
<b>Reviewer Request Associated with "Not Useful" Review</b>			
Use a different model/use a more complex method	30%	48%	.63
Alter the thrust of the paper	27	46	.59

tion, almost 49% indicated that they received a request to provide more complex analyses at least 50% of the time over the past five years, while only 22% indicated they had never received such a request. In contrast, when the authors were talking about useful reviews, 45% of those who were asked to increase the managerial relevance indicated that this resulted in new insights, and 41% indicated it improved the focus of the paper. We take this as additional evidence that our field believes the review process is overemphasizing complexity because, in general, successful authors receiving these reviews do not believe such requests enhance the paper. (We can only imagine the reactions of unsuccessful authors.)

A separate study of 212 authors who had published in *Marketing Science* during the past ten years further addressed these issues. Specifically, it asked the two following questions: (1) "One comment often heard is that the review team is too 'involved' in the review process and asks the author(s) to rewrite the paper in a different direction than the initial submission. Do you think this happens at *Marketing Science*?" and (2) "Another comment heard is that the review team asks the author(s) for more complex analysis/models even though such analysis/models are not necessarily needed. Do you think this occurs at *Marketing Science*?" Approximately 80% of respondents indicated that both requests occurred most or all of the time.<sup>2</sup> They were

<sup>2</sup>Exact responses were as follows: Does *Marketing Science* ask authors to rewrite the paper in a different direction? 18% of respondents answered "all of the time," and 60% answered "most of the time." Does *Marketing Science* ask for more complex analyses? 29% of respondents answered "all of the time," and 45% answered "most of the time."

also asked that if such requests occurred, should the editor step in and say that these requests are not the prerogative of the review team. Approximately 80% agreed the editor should take such an action. (We return to this recommended editor action later in this essay.)

Because the four journals we focused on differ in focus, we analyzed the respondents' reactions by journal. The results (Table 2) are noteworthy: While there is a clear endogeneity issue here (i.e., authors select journals partly on the basis of the type of reviews they expect to get), the journals indeed have different profiles.

The first row of Table 2 indicates that, for approximately 20% of the respondents, *JM* is the journal they read most and in which they publish most often (and incidentally is the most highly cited of the four), while *JCR* accounted for 40% of the respondents. The other two journals each accounted for approximately 20%. (The remaining 3% of the respondents indicated that their preferred journal was not one of the four listed journals.) The second row shows that the four studied journals attract similar mixes of new and seasoned authors, with the length of time our respondents have submitting papers to these journals being approximately 12 years. The next six rows of Table 2 summarize authors' evaluations of the journal that they publish in or read most frequently. While all authors report that the journal's published articles tend to use overly complex statistical methods, are often written too technically, and use too academic a tone, *JMR* and *Marketing Science* tend to be especially high on those dimensions. Despite this, 65% of the authors report that their papers improved as a result of the review process most of the time, and 43% report the review process was satisfactory most of the time. This sug-

**TABLE 2**  
**Authors' Evaluations of the Journal in Which They Publish Most**

Measures	<i>JMR</i>	<i>JCR</i>	<i>JM</i>	<i>Marketing Science</i>
% of authors reporting this journal as the one which they publish in or read most frequently	16%	40%	20%	21%
Authors' length of time in field (1 = "2 or fewer years," and 5 = "more than 20 years")	3.52	3.48	3.48	3.67
<b>Papers in This Journal</b>				
Emphasis on theory (1 = "too little," and 5 = "too great")	2.86	3.39	2.84	2.80
Use of statistical methods (1 = "too elementary," and 5 = "too complex")	3.52	2.98	3.30	3.72
Writing (1 = "too elementary," and 5 = "too technical")	3.48	3.18	3.14	3.59
Tone (1 = "too practical," and 5 = "too academic")	3.48	3.18	3.14	3.59
My papers were improved by reviews (1 = "never," and 5 = "always")	3.45	3.62	3.86	3.63
The review process was satisfactory (1 = "never," and 5 = "always")	3.29	3.20	3.50	3.43
% of authors reporting this journal to have provided the "useful" review they evaluated	29%	26%	26%	12%
% of authors reporting this journal to have provided the "not useful" review they evaluated	27%	30%	12%	19%

Notes: Based on N = 249 advertisements. Attention duration for the advertisement as a whole is not the exact sum of attention to the separate ad elements, because attention to overlapping ad elements is assigned to both. Raw JPEG file size is shown; log-file size is used in the analyses.

gests the review process needs revision, but not necessarily a complete overhaul.

To examine whether expressed concerns with the review process differed by years in the field (in particular, to determine whether the concerns were more strongly expressed by senior colleagues), we analyzed the responses using “time in the field” dummy variables (representing four cohorts: in the field 5 or fewer years, 6–10 years, 10–20 years, and more than 20 years). For most of the items in the survey, there were no significant effects for different cohorts. Thus, the concerns appear to be general. However, there were two questions for which there were cohort differences. The newest cohort (in the field 5 or fewer years) reported lower levels than other respondents for “more complex analyses were required” and “the frequency with which reviewers made conflicting requests.” Receiving fewer requests for more complex analysis is consistent with the newest faculty in the field, who prefer more complex methods. Perhaps younger scholars get fewer conflicting requests because their papers are more focused and thus there are fewer directions/problems that the review team might comment on or because new faculty using these more complex approaches tend to draw reviewers who not only can evaluate but also share in their preference for the new approaches.

We also examined the open-ended comments respondents made (for a sample of such comments, see the Appendix). While this analysis is less formal, many of the responses reinforce the conclusions based on the structured questions. Importantly, none of the responses indicated that the field lacked technical sophistication or that the executional quality of the papers published in marketing was inadequate.

Another way to assess the relative weight that should be placed on different aspects of a paper is to consider how the field ultimately evaluates published papers. We did this by examining the papers that won one of the 40 major research paper awards from *JM*, *JMR*, *JCR*, and *Marketing Science* between 2005 and 2009. Because slightly more than 1220 papers were published in these journals in that time frame, the probability of winning one of these awards is approximately 3.2%. Of the 40 paper awards, 9 (22.5%) were given to papers that had early involvement with the Marketing Science Institute (MSI), yet MSI only published a total of 108 working papers during that period. Even if we assume that all 108 MSI-related papers were published in these journals (which they were not), these data imply that the MSI-related papers were at least three times as likely to win an award as a random paper, suggesting that papers based on “big ideas,” or at least ideas that interest marketing practitioners (and thus are relevant), are actually likely to be highly valued by the field, if they can get published. This observation also supports the conjecture that it is often best to start with a problem and then, and only then, determine the data and methodology needed to address it. Furthermore, if we and others are right that the review process is slanted against ideas and toward execution, the percentage of MSI-involved papers is even more surprising because one might hypothesize that these papers should have had

more difficulty in the review process and therefore would be less likely to get published.

## Promotion and Tenure Decisions

Another major impact on the type of research that is valued is the process associated with promotion and tenure. We examine the promotion process and the criteria that many of our institutions use to show how this process affects the emphasis on rigor and the definition of “sophisticated” research.

A major change in the promotion and tenure process over the past decade is the ability to quickly and inexpensively assess citation frequency. Citation counts have always provided one (but only one) imperfect measure of an article’s impact. However, given the increased ease of data collection, citations have risen in importance and thus are widely discussed in tenure and promotion cases. Aside from arguments about the appropriate source of citation counts (ISI’s Web of Science focuses only on published articles and, in general, is preferred to Google Scholar, which includes working papers that have not yet passed the screen of peer review), several other issues exist. The first is the breadth of a person’s research agenda as well as the relevance of this agenda to the (much larger) basic disciplines. Generating a paper that is cited outside marketing can lead to more citations and, thus, an improved chance for tenure and for promotion. At one level, having impact outside marketing is very desirable because not only does this broaden the reach of marketing, it also generally requires the researcher to address issues believed to be important to the basic discipline. However, obtaining this reach can have the unintended side effect of leading the author to adopt the “sophisticated” conventions of the outside field. This is particularly true in cases in which the basic discipline has embraced “high-tech” methodologies and/or formulaic methodological approaches.

A second issue associated with citation counts is that although they have the advantage of being objective and favoring broadly relevant papers, they are also weighted toward topics that were relevant in the past and favor certain types of papers (e.g., reviews and new methods papers) over others. They also only capture direct references. A possible way to get around this latter point is to develop the entire chain of citations that evolve from a paper, similar to the approach that Toubia (2006) suggests in his article on idea generation. Finally, we note that citation counts are subject to game playing (e.g., citing authors you think, or hope, will review a paper).

Of course citation counts are not the only measure used for promotions. More broadly, many tenure letters and tenure cases turn, for better or worse, on counts of “A” journal publications. This discourages junior authors from putting too much in one paper; encouraging them instead to split a really good idea into its various components to maximize potential publications. A consequence of this approach is that to get the paper’s contribution above some perceived threshold for publication, the author loads this “smaller” paper with methodological bells and whistles. This increases the paper’s rating on execution, and thus probab-

ity of acceptance, but it also makes the work more complicated and less impactful overall.

## PhD Programs

The changing landscape in terms of the review process and the subtle changes in the promotion process have not only altered the field's research agenda but also have affected marketing PhD education, which, as we show in Figure 1, has major consequences for the conduct of research and reviewing, creating a (vicious?) cycle driving research in the field. A few decades ago, most PhD programs included substantial training in relevant marketing institutional knowledge. However, this emphasis has changed. For example, at one time, a major concern in business-to-business marketing PhD programs was to increase rigor as well as relevant knowledge (Danneels and Lilien 1998). Thus, rather than exposing students to a broad overview of marketing, PhD programs have become more specialized, focusing on the technical details of a subspecialty's methodological approaches. Such training not only appropriately encourages the student to acquire knowledge in the basic disciplines but also subtly signals that these disciplines, and the practices and methods they employ, are superior to those in marketing and business. To enhance the success of such specialized programs, PhD students are increasingly drawn from underlying disciplines (e.g., psychology, mathematics, statistics). Such students frequently come into a PhD program with little or no marketing or business experience. Because many of these students are no longer required to take qualifying exams in all areas of marketing (let alone business),<sup>3</sup> they do not develop a sense of the range and complexity of marketing/business problems. Coupling students' lack of exposure to important marketing problems with the pressure they face to immediately begin a research program, these technically trained students tend to extend some previously published idea using their newly acquired technical expertise (often working on a project initiated by their mentor).

The emphasis on depth versus breadth has also been reinforced in the recruiting process. Many leading schools now look for, and hire, the most technically sophisticated recent graduates to "upgrade" their faculty with modern tools and thought. Less emphasis is placed on the quality of the ideas associated with a person's thesis and more on the person's technical sophistication and knowledge, whether measured in terms of some new theory in psychology or sociology or some new method of statistical estimation or economic analysis.

## Suggestions for the Future

Although the preceding discussion points to several issues that merit serious consideration, we believe there is great potential for further advancement of knowledge in our profession. Thus, in the spirit of being proactive, we discuss some potential options to help rebalance idea quality and

<sup>3</sup>Two of the authors took qualifying exams in such subjects as finance, accounting, and management in addition to their major qualifier.

execution and, in the process, redefine what it means to produce a sophisticated research paper. We do this by first focusing on the review process and then provide some suggestions for authors.

### *What Editors and Reviewers Can Do*

In his article, Ellison (2002) suggests that journal editors can act as forces to rebalance the review criteria. By accepting papers that reviewers have rejected because the execution of the paper is not "sophisticated" even though the paper addresses an interesting problem, the editor can provide reviewers with new information about the appropriate balance between idea quality and methodology. Ellison points out, though, that editors are reluctant to overrule reviewers because they rely on them to evaluate papers and do not want to be perceived as "lowering standards." One way to reduce the emphasis on methodology is to institute the criterion of "not being wrong." Thus, reviewers would be asked to comment on the paper's methodology in terms of whether the paper was wrong and, if so, how a different (and presumably a correct) methodology would yield a different result. Thus, for example, it would not be enough to say that a researcher had omitted a factor in the analysis; it would also be incumbent on the reviewer to indicate how this omitted variable would be expected to interact with the variables of interest and alter the general findings.

We believe the appropriate technical concern when reviewing a paper is whether the approach renders the results wrong, not whether the approach is the most advanced approach available. Yes, conceptually, five measures are "better" than one or even four, but does the fact that the researcher only used one measure mean that the results should be ignored?<sup>4</sup> Likewise, longitudinal studies can control for common method bias, but there are times when cross-sectional studies may be acceptable (see, e.g., Rindfleisch et al. 2008) or necessary (e.g., when time-series data are not available). More broadly, a more complex method should not be required unless the results would be materially changed if that more complex approach were used. Such a criterion was used when *Marketing Letters* was founded. Indeed, the only two specific questions asked of reviewers were, "Is the paper interesting?" and "Is the paper not wrong?" These criteria of being interesting and not being wrong were also in place at *Marketing Science* in the mid-1990s (Staelin 1995). It is time for the field to more universally embrace these criteria and ask reviewers to focus on whether the paper meets the two criteria: (1) interesting and (2) not wrong. Of course, the evaluation process should also include a section on how to improve the exposition, interpretation, and so forth, but the principal emphasis should be on helping the authors better explicate their work, not specifying that the analyses need to be reformulated or requiring that the authors use a more "sophisticated" approach simply because they (or at least the reviewer) can. Recalling the survey responses we reported previously, there is widespread support in the field for such an adjust-

<sup>4</sup>See, for example, Bijmolt and Pieters (2001), who show that using single measures often yield qualitatively the same results.

ment to review criteria. Approximately 80% of respondents indicated that the editor should tell reviewers that such requests are not the prerogative of the review team.

Early on, we noted that good research needs good methodology. What then is our definition of a sophisticated research paper? We put forth four criteria: A paper should be (1) reasonably realistic/general, (2) relatively simple and robust, (3) insightful (meaningful), and (4) reasonably communicable. We then apply these criteria to two broad sets of research papers currently found in our literature.

*Behavioral papers.* The current approach for most behavioral work can be caricatured as (1) find some theory; (2) run a series of lab experiments that examine the theory and identify an interaction (i.e., a moderating effect), preferably one that reverses the original result; and (3) report F and *p*-values for only those manipulations and experiments that “worked.” Main effects are considered uninteresting/inadequate for publication (so much for  $F = ma$ ,  $E = mc^2$ ). Cute manipulations, such as a short paragraph used to induce a (temporary) emotion, are often viewed as clever and sophisticated.

We acknowledge that many papers of this type have produced substantial contributions (i.e., the paper is insightful). However, by forcing papers to fit into this paradigm, many issues tend to be ignored. These include general concerns about external validity (i.e., is the paper reasonably realistic?). There is also a question about the duration of effects: Would the manipulation still have impact an hour later or even after it was immediately followed by the reading of a different paragraph? (i.e., are the results reasonably general?) Moreover, by starting with an existing theory and not a problem or a new theory, there is less opportunity for a big idea. Instead, the emphasis is on execution. Why not think of sophistication in terms of a “neat” idea (meaningful/insightful) and not the application of standard procedures associated with the classic behavioral approach?

*“Quant” papers.* As another example, consider the current popularity of structural models and empirical industrial organization. This methodological approach is based on the obviously correct observation that individual actors (e.g., consumers, companies) both anticipate and react to the actions of other parties. It follows logically that failure to account for these actions may (note the word “may” and not “will”) lead to results (e.g., parameter estimates) that are misleading/biased/wrong. In essence, this style of research is the logical descendent of the simultaneous equation regression modeling of the 1970s, with added emphasis on strategic behavior inspired by game theory and analytic modeling.

The beauty of this approach is that it acknowledges possible reactions. However, it also requires the researcher to precisely state the strategic behavior of all the players (i.e., what they know, what they pay attention to, and how far in the future they consider and plan). Moreover, the parameter estimates (and even the ability to identify some parameters) rest on these underlying assumptions. If these assumptions are incorrect, the resultant parameter estimates are biased. What is normally not known is the degree to which this bias exists (i.e., a violation our robustness criterion). This is of

particular concern if the intended use of the estimates is input into a policy simulation to determine the optimal response of one of the players conditional on the reaction of the other players.

Such problems with “quant” papers lead us to believe that “sophisticated” use of such tools should rely more on examining the reasonableness of the underlying assumptions (e.g., by examining the behavioral literature) and/or doing “assumption sensitivity analysis” than on complex estimation methods. Simply asserting that people are completely knowledgeable, have infinite time horizons, and are “rational” seems more naive than sophisticated. Another approach would have the researcher also present basic descriptive statistics and reduced-form ordinary least squares (OLS) regression results. Comparison with OLS results would enable the reader to assess the degree to which the underlying assumptions are driving the solution (thereby addressing the issue of robustness). As a side benefit, because basic statistics and OLS results are “simple,” they enhance the communication of the results.

### **What Authors Can Do**

We are encouraged (see the Appendix) that the overall sentiment among authors (and, by inference, reviewers) tilts slightly toward more emphasis on interest and relevance and less toward complexity for its own sake. Nonetheless, the long-term trend has been in the opposite direction, and a “tipping point” may be rapidly approaching. As published work becomes more specialized and arcane, more potential readers both in business and academia will start to ignore it and turn to other sources for information and intellectual stimulation. In turn, the absence of their influence may encourage researchers to become even more specialized and arcane. This can lead to ever-smaller groups of researchers who reinforce their approach and even codify it as dogma. The end result of this is not desirable, because in the limit each researcher would end up talking only to him- or herself.

Fortunately, this end result is not inevitable. Faculty members in their role as authors (and peer reviewers) have the power to change the system in many ways. As an example, we supplement Ellison’s (2002) prescription for change with a potential solution to the simple versus complex approach debate: When presenting empirical work, always report the results of a simple approach (e.g., OLS regression). This invites comparison to the more complex/sophisticated approach and the application of Occam’s razor—that is, if two approaches generate the same result, the simpler one is preferred. It imposes an implicit requirement that the more complex approach be worth its complication in terms of producing a different interpretation rather than just, for example, a significant but small improvement in the log-likelihood function. Furthermore, publishing the simple approach encourages knowledge accumulation through meta-analysis. (In this regard, a useful requirement of empirical work would be to report means of and correlations among all key variables. This is easy to do now given that the journals also now publish electronic versions, which can contain large appendixes, tables, and so forth.) Finally, provided that the results are qualitatively similar, it

would increase the communicability of the results and thus their potential to be understood at least on a basic level by a broader audience (including sophisticated but busy academic readers, MBA students, and technically sophisticated managers).

In analytical work, the gold standard is a tight closed-form solution. This does well on the simple criterion, but often at the expense of being realistic. As in the case of empirical industrial organization, many of the assumptions sacrifice “reality” to find a closed form solution. In decades past, numerical methods, grid searches, and so on, were not practical to apply. Currently, however, such methods are widely available and extensively used in, among other areas, the basic sciences. These allow for analyzing a more complex model with more believable assumptions, in effect (and paradoxically) meaning that an “inelegant” numerical method can examine more complex phenomena.

An issue associated with numerical approaches is how to display the results, because often the solution changes noticeably over the range of parameters examined. Here, presenting representative cases is useful. Even better, it is usually feasible to show (e.g., via regression) how the key dependent variable depends on the various levels of the model parameters, thus providing a (stochastic) closed-form solution. The resulting “solution” may even inspire or suggest a “law” that can subsequently be proved. (It is a lot easier to “prove a result” when you know what it is in advance).

### **What Faculty Can Do**

Just as we advocate balance in the reviewing criteria, we also advocate balance in PhD education and in hiring and promotion decisions. Both methodologically oriented and substantively oriented marketing academics are valuable, and both should be nurtured. This leads us to the following possible strategies for promoting a better balance between rigor and relevance.

#### *PhD Training and Recruitment*

- Give weight to the applicant’s experience or at least interest level in marketing related issues in the admissions process.
- Require PhD students take an MBA-level marketing class and/or have knowledge of a broad range marketing issues before leaving the PhD program. (Note this should also help them when they begin teaching.)
- Teach doctoral seminars on substantive marketing issues.
- Encourage PhD students to start a research project with an emphasis on a problem versus starting with a data set or a method.
- Write papers and discuss with PhD students why blindly following reviewers’ comments is not always the best course of action.

#### *Faculty Hiring and Promotion*

- Hire and promote marketing faculty who tackle major marketing issues.
- Measure the impact of a faculty member’s research with multiple criteria that include the extent to which the work has changed the way marketing managers think about the issues and/or practice marketing.

- Read and evaluate faculty research instead of relying on the status of the journal in which the research was published. More generally, rely less on whether it is published in an “A” journal and more on the extent to which the research provides new insights.

#### *Research Presentations and Seminars*

- Reserve a substantial fraction of research presentations and seminars for topics that are broadly relevant to different types of marketing scholars as well as scholars in related fields.
- Discourage “gotcha” questions.
- Remember to comment on what is good/interesting rather than just what is imperfect or what you would have done had you chosen to work on the topic.
- Get to the meat of the research (e.g., the results) rather than getting bogged down in discussions of what others have written (i.e. the literature review) or the hypotheses.
- Remember it is their paper, not yours.

### **Summary**

Given our discussion, the reader might wonder when the use of more complex approaches is ever appropriate. The answer is that more complex approaches are appropriate when at least most of the following criteria are met:

1. The assumptions of the more complex approach are met (and hopefully supported logically and/or empirically).
2. When there is strong theoretical support or accumulated evidence for its superiority (e.g., controlling for firm-specific effects in cross-sectional, time-series analysis of firm-level data, accounting for random individual differences).
3. When the more complex method is the focus of the research.
4. When using the more complex approach materially changes the result/interpretation.
5. When, for the intended evidence, the extra communication cost is less than the benefit provided.
6. When the method is described as simply as possible.

The major tenet of this article is that sophistication does not mean complicated. From Occam’s razor (the simpler solution is preferred) to pure math (in which elegant solutions are simple vs. brute force), sophistication is associated with simplicity. The same is true for many schools of design and style, in which simple “clean lines” and limited colors (e.g., black and white) are considered elegant and sophisticated.

In the marketing literature, many of the most influential and most highly cited papers are, at their core, simple. For example, the Bass (1969) model begins with a differential equation but ends up with essentially a quadratic model that nests various diffusion patterns. Guadagni and Little’s (1983) work captures state dependence simply by including an individual-level brand loyalty variable based on the last brand purchased. Thaler’s (1985) transaction utility model has only two terms.

Most of the key behavioral papers examining context effects are simple in concept—for example, Huber and Puto’s (1983) attraction effect and Simonson and Tversky’s (1992) extremism aversion. Even Kahneman and Tversky’s (1979) prospect theory, while complex in math, is basically a simple concept. It is also important that none of these papers apply complex statistical procedures to support their

premises. Interestingly, Wubben and Wagenheim (2008) report that for many purposes, managerial heuristics perform at least as well as more elegant approaches that use Pareto/NBD or BG/NBD models.

Rather than “form follows function” (Mies van der Rohe), it seems at times that the mantra is “I have a hammer, so I’ll hammer in the morning (and afternoon and evening) at pretty much anything I can reach.” There is nothing wrong with developing and trying out new approaches; indeed, this is one way science advances. Still, the fads and seemingly indiscriminate use of new tools after they are developed (without demonstrating that they produce noticeably different and better results in the specific application at hand) have been a bit overdone (e.g., in the past, LISREL/PLS/EQS and, more recently, hierarchical linear modeling, hierarchical Bayes, and hidden Markov). Similarly, topic areas and theoretical perspectives follow predictable patterns (e.g., regulatory focus theory, fluency, construal, social networks). Here again, such activity often leads to important findings, but these mass stampedes are not exactly elegant or sophisticated and predictably produce increasingly incremental insights even when the papers appear to be increasingly sophisticated (or at least complex).

Consider the use of hierarchical Bayesian methods for capturing heterogeneity. Is it really more sophisticated to assume that everyone comes from some distribution with “diffuse priors” (i.e., that I know nothing about the individuals) than to assume (based on theory or prior work) that individuals come from  $K$  groups with known properties (i.e., a latent class approach)? Many would answer “yes” and argue that it is not possible to precisely define the groups in advance. Interestingly, if it is never appropriate to use prespecified groups, why is it appropriate, and sophisticated, to assume that we know exactly how consumers and competitors behave and “constrain” a model in that way? Our point is not that one approach is right and the other wrong. Rather, the point is that choices such as these involve trade-offs that are rarely considered by devotees of one approach or the other. Sophisticated research is something of a craft, which can be approached with multiple tools, including (and quite effectively in many cases) with simple tools such as means, correlations, and regression.

Many years ago, one of the authors had a degeneracy problem in trying to get a multidimensional scaling configuration out of ratings of relationships among foods, nutrients, and benefits to the body. Because this author was attending a seminar by one of the developers of the KYST (Kruskal, Young, Shepard, and Torgerson) scaling algorithm, he asked for help. The expert politely put the output aside, asked what the variables were, and then asked for the raw data. After circling some of the numbers and asking what they were, he looked up and said, “You have the four basic food groups. Why do you need to scale them?” It was a really good question.

The moral of this story, and this article in general, is that complicated is not necessarily sophisticated, and sometimes it is just, well, complicated. Sophisticated research involves good taste in problem selection (i.e., studying something at least that some other relevant or important people would care about), adaptability in analysis, the use of multiple

methods and sensitivity analysis to ensure generalizability, and the willingness to reconsider raw data and/or basic assumptions. It involves nuanced understanding at least as much as cutting-edge methods. In short, sophistication and elegance require uncommon sense. Hopefully, our field will reward and see more of this in the future.

## Appendix

We also examined the comments provided by the respondents. Of the ones that appeared multiple times, several were predictable. These included frustration with conflicting reviews, multiple rounds with changing requests, what are perceived to be technically incorrect reviews, and unblind reviews and favoritism. More relevant here, and consistent with our thesis, were the following comments:

- “The desire for explicit evidence of ‘theory’ and ‘psychological process’ in behavioral papers has gotten really out of hand—leading to narrow and uninteresting work getting published.”
- “Most of the articles in JCR are for psychologists—they are filled with way too much theory and not enough applications. JMR is way too analytical with little relevance to actual marketers. Is it any wonder that the highest impact marketing journal is JM?”
- “Reviewers in Psych are much more likely to accept approaches that are acceptable but are not their favorite.”
- “Psych reviewers understand that ideas have a shelf time and that there is no perfect research (if the idea is interesting).”
- “The Area Editors and most of the reviewers are locked into the classic Churchill ‘multiple items are always necessary’ notion and they focus on statistical assessment of measures.”
- “Journals are unbalanced—i.e., emphasizing only theory instead of both managerial insight and theoretical insight.”
- “The review process sharpens ideas and imposes discipline, but it also has the risk of making things less interesting. The perfect is the enemy of the good.”
- “Too much focus on ‘building theory’ and complex interactions, at the detriment of having managerial impact. I honestly think that as a field, we are actively making ourselves obsolete in terms of usefulness to the profession.”
- “The discipline (and therefore virtually all the journals) are far too focused on methodology at the expense of substantive insights. Publishing marketing research appears to be more about ‘tweaking’ (and unnecessarily complicating) methods than it is about generating useful insights for managers.”
- “Not enough focus on relevancy. Conversely, and here I speak more as a reviewer, there’s a predilection for complexity where simplicity would serve everyone better.”
- “There’s a general tendency among reviewers (and some editors) to look for complexity for the sake of complexity. This is unfortunate and doesn’t always lead to improved research outcomes.”
- “There’s not enough emphasis on (a) interestingness of the research question and (b) robustness of basic result (replication within paper). Instead there is too much emphasis on unique interpretation. There’s also not enough ‘big picture’ view in reviewing.”
- “I don’t see the value of forcing researchers to use a more complicated method of analysis if that method wouldn’t change the implications of the paper.”

---

## REFERENCES

- Bass, Frank M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science*, 15 (January), 215–27.
- (1995), "Empirical Generalizations and Marketing Science: A Personal View," *Marketing Science*, 14 (3, Part 2 of 2), G6–G19.
- Bergquist, Lars and John Rossiter (2007), "The Predictive Validity of Multi-Item Versus Single-Item Measures of the Same Constructs," *Journal of Marketing Research*, 44 (May), 175–84.
- Bijmolt, Tammo H.A. and Rik G.M. Pieters (2001), "Meta-Analysis in Marketing When Studies Contain Multiple Measurements," *Marketing Letters*, 12 (May), 157–70.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), "A Theory of the Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100 (5), 992–1026.
- Blattberg, Robert and Ken Wisniewski (1989), "Price-Induced Patterns of Competition," *Marketing Science*, 4 (Fall), 291–309.
- Danneels, Erwin and Gary L. Lilien (1998), "Doctoral Programs in Business-to-Business Marketing: Status and Prospects," *Journal of Business-to-Business Marketing*, 5 (1), 7–34.
- Ehrenberg, A.S.C. (1995), "Empirical Generalizations: Theory and Method," *Marketing Science*, 14 (3, Part 2 of 2), G20–G28.
- Ellison, G. (2002), "Evolving Standards for Academic Publishing: A q-r Theory," *Journal of Political Economy*, 110 (5), 994–1034.
- Fitzsimons, Gavan J. and Vicki G. Morwitz (1996), "The Effect of Measuring Intent on Brand-Level Purchase Behavior," *Journal of Consumer Research*, 23 (June), 1–11.
- Guadagni, Peter M. and John D.C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (3), 203–238.
- Huber, Joel and Christopher P. Puto (1983), "Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects," *Journal of Consumer Research*, 10 (1), 31–44.
- Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–91.
- Lehmann, Donald R. (2005), "Journal Evolution and the Development of Marketing," *Journal of Public Policy & Marketing*, 24 (Spring), 137–42.
- Lodish, Leonard (1986), *The Advertising and Promotion Challenge: Vaguely Right or Precisely Wrong?* Oxford: Oxford University Press.
- McAlister, Leigh (2005), "Unleashing Potential," *Journal of Marketing*, 69 (October), 16–17.
- Reibstein, David J., George Day, and Jerry Wind (2009), "Is Marketing Academia Losing Its Way?" *Journal of Marketing*, 73 (July), 1–3.
- Rindfleisch, Aric, Alan J. Malter, Shankar Ganesan, and Christine Moorman (2008), "Cross-Sectional Versus Longitudinal Survey Research: Concepts, Findings, and Guidelines," *Journal of Marketing Research*, 4 (August), 261–79.
- Simonson, Itamar and Amos Tversky (1992), "Choice in Context: Trade-Off Contrast and Extremeness Aversion," *Journal of Marketing Research*, 29 (August), 281–95.
- Staelin, Richard (1995), "Editorial," *Marketing Science*, 14 (1), 1–4.
- Swanson, Edward P. (2004), "Publishing in the Majors: A Comparison of Accounting, Finance, Management and Marketing," *Contemporary Accounting Research*, 21 (1), 223–55.
- Thaler, Richard H. (1985), "Mental Accounting and Consumer Choice," *Marketing Science*, 4 (3), 199–214.
- Toubia, Olivier (2006), "Idea Generation, Creativity, and Incentives," *Marketing Science*, 25 (September/October), 411–25.
- Wilkie, William L. and Elizabeth S. Moore (2003), "Scholarly Research in Marketing: Exploring the '4 Eras' of Thought Development," *Journal of Public Policy & Marketing*, 22 (Fall), 116–46.
- Wubben, Markus and Florian v. Wagenheim (2008), "Instant Customer Base Analysis: Managerial Heuristics Often 'Get It Right,'" *Journal of Marketing*, 72 (May), 82–93.
- Yadav, Manjit S. (2010), "The Decline of Conceptual Articles and Implications for Knowledge Development," *Journal of Marketing*, 74 (January), 1–19.