

Knowledge Creation in Consumer Research: Multiple Routes, Multiple Criteria

John G. Lynch, Jr.
University of Colorado

Joseph W. Alba
University of Florida

Aradhna Krishna
University of Michigan

Vicki G. Morwitz
New York University

Zeynep Gürhan-Canli
Koç University

John G. Lynch, Jr. is the Ted Andersen Professor of Free Enterprise at the Leeds School of Business at the University of Colorado. Address: 419 UCB, Boulder, CO 80309-0419 (Phone: 303.492.8413; Fax: 303.492.5962; email: john.g.lynch@colorado.edu). Joseph W. Alba is the James W. Walter Eminent Scholar at the Warrington College of Business Administration at the University of Florida. Address: 212 Bryan Hall, P.O. Box 117155, Gainesville, FL 32611-7155 (Phone 352.377.8213; Fax: 352.846.0457; email: joe.alba@cba.ufl.edu). Aradhna Krishna is the Dwight F. Benton Professor of Marketing at the Ross School of Business at the University of Michigan. Address: 701 Tappan Street, Ann Arbor, MI 48109 (Phone: 734.764.2322; Fax: 734.936.8715; email: aradhna@umich.edu). Vicki G. Morwitz is the Harvey Golub Professor of Business Leadership at the Stern School of Business, New York University. Address: 40 West Fourth Street, Suite 807, New York, NY 10012 (Phone: 212.998.0518, Fax: 212.995.4006, email: vmorwitz@stern.nyu.edu). Zeynep Gürhan-Canli is the Migros Professor of Marketing at Koç University. Address: Rumeli Feneri Yolu, Sariyer, Istanbul, Turkey (Phone: 212.338.1784; Fax: 212.338.1642, email: zcanli@ku.edu.tr).

This is a preprint version of the article. The final version may be found at <https://doi.org/10.1016/j.jcps.2012.06.004>

Abstract

The modal scientific approach in consumer research is to deduce hypotheses from existing theory about relationships between theoretic constructs, test those relationships experimentally, and then show “process” evidence via moderation and mediation. This approach has its advantages, but other styles of research also have much to offer. We distinguish among alternative research styles in terms of their philosophical orientation (theory-driven vs. phenomenon-driven) and their intended contribution (understanding a substantive phenomenon vs. building or expanding theory). Our basic premise is that authors who deviate from the dominant paradigm are hindered by reviewers who apply an unvarying set of evaluative criteria. We discuss the merits of different styles of research and suggest appropriate evaluative criteria for each.

For nearly as long as its existence, consumer research has wrestled with questions about its mission, direction, boundaries, and relevance (Belk, 1984; Deighton, 2007; Folkes, 2002; Hirschman, 1986; Holbrook, 1985; Jacoby, 1976; MacInnis and Folkes, 2010; Sheth, 1982; Shimp 1994; Simonson et al., 2001; Wells, 1993). More recently we have also witnessed particular concern over the perceived dearth of novel, far-reaching consumer insights (Lehmann, McAlister, & Staelin, 2011; MacInnis, 2011; Reibstein, Day, & Wind, 2009). We do not hope to resolve this dilemma. We do however wish to address a perceived malaise stemming from questions about the nature of knowledge discovery. Such an objective inevitably invokes a philosophical discussion about the rules of science, but we are primarily concerned with the more mundane rules of journals.

Our fundamental premise is that our discipline, although nominally pluralistic about avenues to knowledge discovery, is insensitive to the demands of different styles of inquiry and inflexibly applies evaluative criteria relevant to the dominant mode of inquiry to other legitimate approaches that are more appropriately judged by different criteria. As a result, individual researchers grow frustrated with the review process and, moreover, scientific progress is inhibited. Our objective, then, is to highlight the multiple paths to knowledge discovery and recommend corresponding evaluative criteria.

Deduction and its discontents

The preponderance of empirical investigation in consumer psychology can be characterized by the hypothetico-deductive approach, a scientific approach that involves using theory to formulate hypotheses that can be falsified with observable data (Calder, Phillips, &

Tybout, 1981; Lynch, 1982; Sternthal, 2010). We contend that the dominant style has become dominant because it is perceived to be more “rigorous” than other approaches. However, analytic imperatives become counterproductive when the value placed on research rigor and sophistication exceeds the value placed on the importance of the research question and the substantive insights provided by that research, or, where “complexity is valued at the expense of relevance” (Lehmann et al., 2011). As the problem pertains to journal policy and the review process, both demands are captured by Ellison’s (2002; see also Rozin, 2009) distinction between q-quality and r-quality, the former “reflecting the importance of a paper’s main contribution” and the latter “reflecting other aspects of quality (generality, robustness checks, extensions, discussions of related literature, etc.” (p. 995). Ellison’s gloomy contention is that, for many years and across a variety of academic disciplines, improvements in r-quality have not been matched by increases in q-quality; indeed, the perceived (and real) importance of r-quality has resulted in a misallocation of researchers’ efforts and, therefore, the improvements in r-quality have come at the expense of q-quality.

We agree, but argue further that the perceived importance of r-quality has not only influenced authors’ efforts, but has had an unfortunate negative impact on the journal review process. As pointed out by Ellison (2002) and Lehmann et al. (2011), reviewers are inclined to focus on technical rigor (the r-quality) at the expense of “importance” or relevance (the q-quality). This tendency is not surprising, inasmuch as it is much easier to obtain inter-reviewer agreement about technical issues, such as the existence of a confound, than about the “importance” of the findings. Assessments of q-quality require subjective judgments of the magnitude of belief shift produced by the research and the importance of that belief shift —

judgments that can be idiosyncratic due to differences across reviewers in terms of their prior knowledge of the domain, appreciation of the implications of the research, or beliefs about the discipline's priorities. Although we discuss possible remedies to the loss of q-quality, we further argue that assessments of r-quality or "rigor" can be biased by methodological orthodoxy and that reviewers can suffer from certain misconceptions even in these assessments.

Our goal therefore is to provide guidance to reviewers for how to evaluate research so that q-quality is not neglected, and r-quality is properly upheld but not overemphasized. Although we provide guidance for traditional hypothetico-deductive research, our main goal is to offer guidance for other paths to knowledge creation, that we think might even offer more promise for higher levels of q-quality, and where we think over-emphasis of and misperceptions concerning r-quality are more problematic. In the next section, we discuss different domains to knowledge creation, and offer guidance for reviewers.

Domains and routes

Our discussion is organized around two fundamental dimensions: the domain to which research is intended to contribute (to building or expanding theory or to understanding a substantive phenomenon) and the approach by which the contribution is made (by starting with theory – a deductive approach, or starting with phenomenon driven observations – a non-deductive approach). The former is well-known to consumer researchers but is formalized in Brinberg and McGrath's (1985) validity network model, which claims that all research involves elements and relations from the conceptual, substantive, and methodological domains. We restrict our discussion to research intended to contribute to either the conceptual or substantive

domains, and we focus exclusively on empirical research.¹ Such efforts comprise the large majority of consumer research and are the source of much consternation in the review process. The second dimension, the researcher's approach to inquiry, has less of a history but has recently become a topic of conversation. For present purposes, it is sufficient to make a basic distinction between deductive and non-deductive approaches. The combinations of researcher intentions and researcher approaches to inquiry produce four paths to knowledge, as shown in Figure 1.

Insert Figure 1 here.

A key motivation for the present paper is our perception that the community of reviewers within consumer research adheres to the view that research should primarily make a “theoretical” contribution, narrowly defined in terms of construct-to-construct links. Research that illuminates links from constructs to phenomena is viewed as merely applied (a view to which we take exception). It is ironic that Brinberg and McGrath associated “theory” with statements about relationships between concepts and important substantive phenomena. Although we may look askance at research that has no greater ambition than to “demonstrate” that a psychological effect is relevant to consumer behavior, “theoretical” contributions can obtain from testing one or more explanations of real-world consumer decisions, and important contributions can be made by

¹ In so doing, we acknowledge the important but relatively infrequent methodological contributions made via the development of new procedural paradigms (e.g., Mouselab) or novel statistical methods which require quite different evaluative criteria. We also acknowledge the importance of non-empirical papers, both conceptual and substantive, whose infrequent appearance in the literature has been lamented and whose evaluative criteria have been addressed (MacInnis, 2011; Yadav, 2010).

conceptualizing some substantive consumer system in terms of constructs that have been investigated in connection with very different phenomena.

In the remainder of this paper, we separately consider the cells of Figure 1, providing examples from the literature and recommending evaluative criteria appropriate for each. Our overriding assertion with regard to evaluation follows Abelson's (1995) position that research is "interesting" if it shifts one's beliefs about something something deemed important.

We go one step further and argue that different varieties of research attempt to shift different types of beliefs. We will distinguish between updating beliefs about: a) construct-to-construct links; b) construct-to-observable links; c) observable-to-observable links; and d) beliefs that some phenomenon "exists" in the world.

Conceptual contributions via deduction

Research from this category constitutes much of the consumer psychology literature. The primary intention of papers that follow the deductive-conceptual route typically is to make claims about construct-to-construct relations rather than claims about the relationship between concepts and some particular substantive system. The defining feature of such research is not the use of experimental methods but rather a distinct lack of emphasis on the substantive domain. The genesis of these papers frequently is the literature rather than life (Lynch 2011a), as exemplified by research on dual process models of persuasion, regulatory focus, fluency, construal level, priming, mindsets, and embodied cognition. Despite the popularity and maturity of this category of research, we feel there are still deficits in the way research in this category is evaluated.

Criteria for reviewing. Research contributions from this category are generally measured by the degree of belief shift induced about construct-to-construct links as well as the reader's assessments of the generality of those construct-to-construct links and the importance of those links in updating or unifying prior theory. We comment more on these points below. We also argue that consumer research that makes a conceptual contribution via deduction should be evaluated on its relevance to a consumer domain.

Belief shift and "rigor." When evaluating a paper's level of contribution, a chief criterion is belief shift about construct-to-construct relations. The degree of belief shift is affected by the effect size and the judged rigor of both the explanation and methods. Reviewers often penalize a paper that has the potential to produce a large belief shift because they judge it to have methodological flaws that leave open alternative interpretations about construct-to-construct links. Park (2012) refers to informative papers for where there is ambiguity about causation as "cute" rather than "beautiful," and he attributes significant value to some of these papers, despite their judged flaws.

However reviewers often state that such papers lack rigor. Unfortunately, "rigor" can be a misunderstood and selectively applied criterion. When reviewers suggest that the paper suffers from a "fatal flaw," they are obliged to offer an alternative explanation for the data. And, indeed, reviewers experience little difficulty in providing such explanations for deeply flawed research. Researchers are more likely to be frustrated with the review process when reviewers offer competing but inferior alternative accounts for the research that the authors earnestly and systematically conducted though they failed to provide an airtight case. No research is perfect and, as neatly articulated by Gordon Bower (2007), "It is a poor psychologist indeed who cannot

think of an alternative explanation for almost any result.” To be damning, however, such alternative explanations should be plausible and parsimonious (Sternthal, Tybout, & Calder, 1987). In other words, they should be able to explain all the authors’ data and not merely part of it. Quite often, reviewers’ alternative explanation can account for only a subset of the data (i.e., the results of some studies but not all). An author’s comprehensive and face-valid account should not be trumped by a set of idiosyncratic explanations. Of course, it is incumbent on the author to proceed earnestly and provide reviewers with little fodder for such “fundamental reviewer errors” (Lynch, 1998).

Less widely recognized is the Bayesian perspective on belief change and rigor, which questions the unique roles of deduction and falsification in theory testing (Brinberg, Lynch, & Sawyer, 1992). According to Bayes theorem, belief in a hypothesis should increase if the hypothesis predicts a particular outcome, the probability of the outcome is greater under the proposed hypothesis than under a different hypothesis, and the outcome is, in fact, observed. Consequently, findings that have multiple interpretations can nonetheless lead to a shift in beliefs about a theoretical relationship and, therefore, reviewers should be sensitive not only to the parsimony of their multiple explanations but also to the relative strength of a single competing explanation. Brinberg et al. showed that “fatal” confounds may not be fatal; it can be logically proper to have more belief updating from a high-powered study with an alternative explanation than from a low-powered study where the finding has no alternative explanation. Most reviewers would be more persuaded by a study where they see no alternatives.

Judged importance. It is necessary but insufficient for a paper to induce belief shift about construct-to-construct links for papers that attempt to make a conceptual contribution using

deduction. The beliefs must also be perceived to be “important,” frequently expressed in terms of the “level of contribution” of the paper. Importance can be attained from overturning prior theoretical accounts of influential findings or from the perception that the construct-to-construct relations are insightful in terms of the breadth and depth of their explanatory power. Thus, whereas Aaker and Lee (2001) made a significant contribution by demonstrating that goals associated with approach and avoidance and self-view can influence persuasion, the hundredth subsequent test of regulatory focus theory in a persuasion context is likely to be greeted with less enthusiasm.

Judged importance often coincides with the novelty of an idea, but the literature shows that novelty is not a necessary condition. Some “conceptual replications” have great heuristic value for consumer researchers who either are unfamiliar with the original source or who had not deeply considered its deductions relevant to consumer behavior. Many such papers tread a fine line between psychology and consumer research, but it is essential that the judged contribution of these papers be based on their potential to illuminate consumer behavior and stimulate additional research beyond what would have been observed in their absence—a criterion that is not easily met. The ability to illuminate is a function of the distance between the original and applied disciplines. Inspiration from a neighboring field is likely to make a smaller contribution than is inspiration from a distant field (e.g., signal detection theory). It is proper for reviewers to judge a paper as not being sufficiently illuminating so long as they appreciate that research can make a contribution even if precedent exists in another discipline.

Theoretical novelty is expected to be high for papers with a hypothetic-deductive approach. Such papers typically introduce a new mediator or moderator of an existing

relationship or process (Colquitt & Zapata-Phelan 2007). The most innovative ones examine a previously unexplored relationship or process and change beliefs about construct-to-construct relationships that had been viewed as important in the literature (e.g., does involvement decrease attitude change?) and/or introduce a new construct.

Losing the consumer. Research that pursues conceptual contributions via deduction typically makes the substantive area of application an afterthought (Brinberg & McGrath 1986). The consumer often becomes a convenient context to study some more general nomological relations. Consumer research has been defined as “the study of people operating in a consumer role involving acquisition, consumption, and disposition of marketplace products, services and experiences” (MacInnis & Folkes, 2010, p. 900). However, some papers possess only a very thin consumer veneer and cannot be readily distinguished from papers that appear in general psychology journals. Authors need to establish and reviewers need to assess the relevance of the findings for increasing our understanding of consumers.

It is noteworthy that the hegemony of deductive conceptual papers, in which construct-to-construct contributions are prized over explanations of real-world phenomena, has been criticized from within the field consumer researchers wish to emulate (Rozin, 2001). Unfortunately, willingness to snub the consumer in consumer research may be proceeding in the wrong direction. Lehmann et al. (2011) noted the trend in the broader field of marketing to place increasing emphasis on technical rigor at the expense of the (substantive) relevance of the findings, which in turn has led to the recruitment of faculty from basic disciplines whose interests are only loosely tied to marketplace phenomena. These concerns could be applied more narrowly to consumer research.

Substantive contributions via deduction

We turn next to papers that also adopt a deductive logic but intend to make their contribution not primarily by illuminating construct-to-construct links but rather through demonstrating and explaining *substantive* phenomena or interventions designed to change or improve some real world behavior. It is noteworthy that, in marketing journals outside the realm of consumer behavior, papers focusing on the substantive domain are disproportionately represented among award winners, are disproportionately cited, and are more likely than papers on construct-to-construct links to be of interest to those in other fields and to the general public (Lynch, 2011a; see also Lehmann et al., 2011). Such “real-world” papers are not the norm within the realm of consumer research. One reason pertains to a taste-related preference for theoretical development at the expense of application that, at its worst, results in findings that have little substantive relevance or lack pragmatic actionability. In addition to taste, however, we suggest real-world papers are often judged on the basis of inappropriate evaluative criteria.

Deductive approaches to understanding substantive phenomena take two forms. In the first, one combines substance and concepts to form a conceptual hypothesis and then finds some appropriate methods to test the hypothesis. Critically, these are hypotheses about phenomena and construct-to-phenomena links rather than construct-to-construct links. The second approach involves the engineering of important interventions that are based on theoretic concepts from previous research, not unlike the process of “intervention falsification” (Calder et al. 1981). We discuss these approaches in turn.

Explaining substantive phenomena using theoretic constructs. In what Brinberg and McGrath (1986) call the “theoretical path,” researchers seek to explain a *phenomenon* in the substantive domain by mapping relevant constructs to it. “Theory” refers to understanding the phenomenon rather than construct-to-construct links.

Many such papers start with observations of everyday behavior or “stylized facts” to be explained, but the contribution derives from successful application of existing theoretic constructs to explain the observations. For instance, Ariely and Levav (2000) started from the everyday observation that when eating in group settings, people tend to choose something different from others at the table. They deduced that, if group variety loomed large as a goal, those ordering last should be least satisfied with their orders. Field experiments provided support for the hypothesis that being in a group leads consumers to balance individual and group related goals.

Other papers attempt to reconcile theoretical conflicts in a literature regarding a substantive phenomenon. Maoz and Tybout (2002) noted conflicting findings in the literature and conflicting theories about whether moderately incongruent brand extensions are evaluated more favorably by consumers than congruent or extremely incongruent extensions. On the one hand, resolving moderately incongruent information can result in task satisfaction which can carry over to how the brand extension is evaluated. On the other hand, consumers might logically be more likely to transfer their positive feelings for the parent brand to the extension the more they perceive a fit between the two products. Maoz and Tybout reconciled these conflicting accounts by positing that involvement moderates the effect of brand extension congruity on evaluation. They argue that the task satisfaction that results from resolving moderately

incongruent information only occurs when task involvement is high; when it is low consumers are likely not motivated enough to resolve the information.

Yet other theoretical path papers set out to test an interesting conceptual question of substantive importance. Thus Krishna (2011) studied consumers buying a product on cause marketing, where part of the proceeds from the sale of the product go to a cause such as AIDS research. She tested whether such purchases can decrease direct donation to the cause. Yoon, Gürhan-Canli, and Schwarz (2006) provided a theory-based explanation for the circumstances under which corporate social responsibility activities may hurt, improve, or not change a company's tarnished image. Papers in this vein are judged as much on the importance of the phenomena they attempt to explain as on the completeness and tightness of the explanation offered, such as with the earliest attempts to understand brand equity (Aaker & Keller, 1990).

Criteria for reviewing. Research contributions for papers of this type should be measured by the degree of belief shift regarding an issue deemed important, where key beliefs link substantive phenomena to concepts. Rigor is strongly influenced by judgments that essential features of the to-be-explained substantive phenomena have been captured. Belief shift is affected by the empirical force of the data in refuting a prior conceptualization of the phenomena, as well as by the readers' priors about those phenomena.

Prior beliefs and perceived surprisingness. Two readers might have different judgments of the importance of an "effect" if they begin with different expectations. Findings from these papers can also be judged obvious if the focus is inappropriately placed on construct-to-construct links rather than the links from constructs to the consumer phenomenon of interest. Consider Nedungadi's (1990) paper on the phenomenon of consideration sets. Nedungadi argued that

choice could be explained not only by relative evaluations of competing options but also by whether an alternative was considered at all. These principles are remarkably powerful but may seem obvious once stated. Nedungadi showed that “awareness” advertising of a less preferred brand in a “niche” subcategory could actually increase the share of the leading brand in the niche category by cuing retrieval of that otherwise inaccessible subcategory and then prompting retrieval of better liked alternatives in that subcategory. Reviewers on this paper, failing to map memory constructs onto the substantive problems of consideration set formation and choice, objected that Nedungadi’s deductions were “obvious” because they followed from any network model of memory. They misjudged the contribution by asking themselves whether the paper caused updating of construct-to-construct links rather than updating of links between constructs and important substantive phenomena.

Familiarity, imagination, and judged “importance.” Just as readers can differ in their prior beliefs, they can differ in their interest in the phenomena under study and in their ability to derive additional interesting implications of the findings. In the case of Nedungadi (1990), the editor and a reviewer differed along these lines. To the benefit of consumer researchers, the editor held sway and the paper has become the most cited paper on consideration sets in the literature.

Clever empirical demonstrations are valuable in and of themselves but also because they can prompt deeper elaboration of the implications of those findings. Nedungadi showed not only that choice is a function of what is—and what is not—remembered; also choice dynamics may be driven more by changes in consideration sets than by changes in relative attitudes, the latter representing the prevailing perspective at the time. Similar points can be made about other

papers whose goal has been to illuminate the concepts relevant to understanding particular phenomena. For example, Shiv and Fedorikhin (1999) tested ideas of emotional factors on self-control that followed closely from previously made arguments regarding visceral influences (Loewenstein, 1996) and “low road” versus “high road” processes (LaDoux, 1996). However, substantial creativity was involved in mapping those concepts to self-control phenomena of substantial interest to consumers in the real world, such as choosing between fruit salad and chocolate cake. The resultant findings could be “obvious” to a reader with deep expertise in these literatures while at the same time be broadly illuminating to others with different backgrounds who not only benefit from the empirical demonstration, but who also are able to appreciate its implications for other settings of interest.

Judgments about “importance” also stem from assessments of the pervasiveness of the substantive phenomenon under investigation and the reader’s ability to see implications beyond what was shown directly. Lynch and Ariely (2000) examined the effects of lowered search costs for price and quality on consumer price sensitivity. They showed it was true that price sensitivity was increased by lowering search costs for price – as many had observed. However, making it easy to compare the wares of two retailers did not increase price sensitivity if the products sold were imperfect substitutes rather than exact duplicates. Readers who experienced similar belief shifts from the empirical findings will naturally vary in their judged “importance” of those belief shifts based on their personal assessments of the pervasiveness, economic, and psychological significance of explaining price sensitivity in online channels.

A note on field studies and use of incentive-compatible procedures. Because papers of this nature are judged on their ability to illuminate a substantive phenomenon, clever field

studies are especially prized and valuable, as are demonstrations of effects that use incentive-compatible procedures. Contrary to the beliefs of many researchers and reviewers, however, the value is not from an advantage in terms of external validity. External validity is about whether some core result varies as a function of some background factors. Unless a background factor is varied, one cannot make any statement about external validity. Researchers and reviewers are wrong to think that field settings are homogeneous; if only a single setting is tested, there is no more reason to expect that a finding from one field setting will generalize to another field setting than that a finding from one particular lab setting will generalize to some particular field setting (Dipboye & Flanagan, 1979; Lynch, 1982, 1999).

Nonetheless, there are two reasons for awarding a bonus to a paper for having included an ingenious field study or one that uses an incentive-compatible “real behavior” procedure. First, the study helps convince readers that concepts have been appropriately mapped to the underlying substantive phenomenon. In addition, readers may be drawn more deeply into the clever “real behavior” study than to its companion studies and, as noted, much of the judged contribution of a paper derives from the reader’s willingness to elaborate on the likely implications of the findings. We therefore suspect that the persuasiveness of Ariely and Levav (2000) was enhanced by field studies of a Chinese restaurant and a brew pub that showed that because of their group interactions, groups of diners and beer drinkers chose more variety in the dishes they ordered than would be if individuals made independent choices. Similarly, Wansink and Van Ittersum’s (2003) field test of the elongation bias (i.e., elongated objects are perceived to be larger than less elongated objects, and therefore people consume more from a short, wide

glass than from a tall, slender glass) appears to be as well known as its laboratory-based predecessor (Raghubir & Krishna, 1999).

Intervention falsification: Theory-based interventions to influence substantive systems. Field experiments occasionally have a goal not of explaining some real-world behavior but of changing real world behavior via a public policy or managerial intervention. These papers have characteristics that Calder, Phillips, and Tybout (1981) associated with “intervention falsification,” wherein constructs tested in a tight, laboratory study are mapped to a noisy real-world environment for the purpose of testing assertions about the observable-to-observable mappings rather than adding any new insight on construct-to-construct relations or attempting to pinpoint the exact reasons why their interventions work.

For example, experimental and behavioral economists have made significant contributions to consumer welfare and public policy by using existing theory on intertemporal choice to devise interventions that increase retirement savings. Thaler and Benartzi (2004) have trade-marked a plan called Save More Tomorrow, which asks people saving X% for retirement to precommit to save X+Y% once given a raise. The authors deduced from laboratory research on intertemporal choice that people might discount the difficulty of making increased contributions at a more distant point in time. They deduced from prior work on the status quo bias and inertia that once a soft precommitment had been made, people would be unlikely to incur even a small cost to undo it if they discovered their budgets were tighter than expected when the raise was received. Results showed a dramatic increase in annual savings rate for retirement (3.5% to 13.6% over the course of 40 months). Madrian and Shea (2001) deduced from a similar literature on the status-quo bias that 401(k) participation would be significantly

higher under automatic enrollment; they used interrupted time series to support that an auto-enrollment intervention worked as expected.

The substantive contributions of these papers are self-evident. In both cases, however, several explanations for why the intervention works are possible (cf. Lynch & Zauberger, 2006). Such interpretive ambiguity is not a criterion that should be used to evaluate papers of this sort. Few—if any—papers in the consumer research literature have produced equally dramatic effects on social welfare and public policy, yet we speculate that any such papers would struggle in our review process because reviewers would either misapply criteria for tests of hypotheses about construct-to-construct links to this kind of work or denigrate such “mere” application of existing theory.

Criteria for reviewing. The beliefs to be altered by theory-based intervention research are not about constructs but about empirical relations between interventions and real-world outcomes. Effect size matters, particularly so in relation to the costs involved in the intervention, because the research aims to improve system well-being (Brinberg & McGrath, 1985). The judged importance of the effects also depends on the prevalence of the phenomenon. In this sense, an intervention that produces large effects on retirement savings in 401-K programs might be judged to be more important than a study that showed small effects on participation in Christmas club accounts. More attention is paid to the construct validity of the dependent variable than to the construct composition of the independent variable.

What matters in intervention falsification is that the overall package works, not which particular aspect of the intervention was the central ingredient. Consequently, it is extremely important that one can attribute the change in behavior observed to the interventions and not to

something apart from the interventions. This is classic “internal” validity in the sense of Shadish, Cook, and Campbell (2002), as distinguished from issues of construct validity (trait and nomological) that are of concern when attempting to explain results by constructs. The constructs are simply the building blocks for creating the intervention. Intervention studies, by definition, are completed in the field and often require grudging cooperation from external partners. It is unreasonable for reviewers to ask for minor revisions of the study in the casual way many reviewers ask for new laboratory studies.

Conceptual contributions via non-deductive routes

A very different epistemological approach begins not with formal theory but with strong hunches or simply with data. In this collection of methods the data precede rather than follow theory. Many social scientists, consumer researchers included, have scorned such papers as “atheoretical.” We recognize the intensely empirical nature of this approach, but we disagree that it has less scientific credibility. We also object to caricatures of this approach, sometimes with the pejorative term of “dust bowl empiricism” (McGrath & Brinberg, 1983), as lacking in real or potential conceptual contribution. Given the binary choice of having a conceptual explanation for an interesting phenomenon versus not having one, we suspect that the former would receive near-universal preference. The true divide between the deductive and non-deductive camps revolves around the need to have a tight conceptual account *now*.

Indeed, Haig (2005) outlines a sequence that moves from theory generation to theory development to theory appraisal, wherein scientific investigation begins with observation and the probing of data for the presence of a surprising phenomenon and is followed by a tentative but

plausible theoretical account of the phenomenon. To the extent that the investigator is successful, the novelty objective is satisfied and a potential path to deeper insight is identified. Note, however, that the process succeeds even if it “fails.” If a phenomenon is uncovered that resists a tidy theoretical explanation, it still endures. Haig argues for the need to differentiate empirical from theoretical progress and, accordingly, to recognize that the discovery of a phenomenon is a notable achievement by itself—so much so that more Nobel prizes have apparently been awarded for the discovery of phenomena than for explanatory theories of those phenomena. It should also be evident that the process is not only at sharp variance with deduction but also falls short of r-quality. The initial phenomenon may have no theoretical foundation or precedent, and a documented explanation may be a distant goal.

At the risk of reinforcing unfortunate stereotypes, non-deductive methods range in terms of their thoughtfulness. Serendipity anchors one end of the continuum, where discoveries are completely unanticipated. Although there is no shortage of scientific breakthroughs traceable to serendipity, we are reluctant to endorse luck as a method. Nor, however, are we inclined to denigrate luck or eschew attempts to capitalize on luck. With regard to the latter, it is important to note that “planned” luck offers not only a route to discovery but also a potentially efficient route when massive probing can be performed at low cost. Consider, for example, the process of “accelerated serendipity” in chemical science wherein a very large number of chemical reactions are tested without any forethought, but with the hope of an interesting outcome (McNally, Prier, & MacMillan, 2011). Consumer researchers may lack such laboratory techniques, but they can avail themselves of data-mining.

Trial and error has similarly been portrayed as a witless way to approach discovery, but the argument can be made that it occurs with surprising frequency and may even be a preferred approach in some contexts. Wasserman and Blumberg (2010) argue people are inclined to attribute to creative genius discoveries that are truly attributable to trial and error. Methods possessing a trial-and-error flavor have recently been touted as the preferred approach in business and policy contexts. Various descriptions include “adaptive trial and error” (Harford, 2011), “measure and react” (Watts, 2011), “build-measure-learn” (Ries, 2011), and “failing by design” (McGrath, 2011), the argument is that the desired end state, i.e., a successful product or policy, is more efficiently achieved through a bottom-up approach when environments are highly complex and unpredictable. Proponents of this approach do not eschew forethought but instead advocate small-bore hunches that can be easily field-tested in both highly scalable and highly constrained samples and that can offer insight even in the case of failure.

Lest one conclude that what is good for application is not good for formal science, there is evidence, at least from the biological realm, that a similar logic is part and parcel of everyday research. Dunbar (1999) reports that unexpected findings are expected and that researchers deliberately design experiments that allow for such findings—or, in his apt phrasing, “the prepared mind favors chance.”

The process observed by Dunbar does not end with the unanticipated finding. Good science requires additional testing and speculation. The first reaction prompted by an unanticipated finding is skepticism regarding method. Thus, methodological explanations are advanced and tested. If the unanticipated finding proves robust, investigators then advance hypotheses and models.

If investigation is prompted not by a serendipitous experimental finding but with casual observation of the world or the musing of an investigator about the relationships among constructs, the process takes on an abductive flavor. The end result is a tentative hypothesis that can serve as a starting point for a more formal theory. Investigators may also engage in formal induction, starting with a set of observations and reasoning upward toward a rule that can be subsequently tested through a falsification logic. Alternatively, one may simply be satisfied with establishing the robustness of a conceptual assertion, providing the discipline with a stylized fact or empirical generalization.

Taken as a class, these informal routes to discovery present a clear contrast to the traditional deductive logic of consumer research. Papers that take these routes also struggle in the review process due to their “atheoretical” nature. We lament this state of affairs, in part because our own research has on various occasions involved a fair amount of stumbling around. For example, Alba and Chattopadhyay (1985) began with a somewhat muddled set of data regarding the effect of brand cues on brand recall. The relationship became clear and easily explained once the data were reanalyzed using gender as a blocking variable. Gender served as a surrogate for expertise but was included in the design as a background variable with no particular hypothesis in mind. The inclusion of background variables may more generally be viewed as a form of prepared serendipity.

The mere-measurement effect (Morwitz, Johnson, & Schmittlein, 1993), in which asking consumers about their purchase intentions has an effect on subsequent purchase, was not a serendipitous finding but one that qualifies as “findings-first” research. After more than a decade of research, a precise account of the effect remains elusive. In this sense it mimics the hugely

influential attraction effect reported by Huber, Payne, and Puto (1982) which, too, has been slow to garner a definitive explanation. Such “atheoretical” findings-first papers can be of exceptional heuristic value in alerting us to topics of true importance (Lynch, 2011a).

Criteria for reviewing. Findings-first papers that bear on construct-to-construct relations are judged based on belief revision of construct-to-construct links. In this sense, they are like deductive concepts-first papers, but they should not be similarly expected to have those links nailed down. Measurement is not supposed to change the phenomenon under study, but Morwitz et al. showed that measuring intentions actually caused people to go out and buy computers and cars. Typically, the finding itself is expected to be inherently interesting in “findings first” research, so one is also updating beliefs about whether some observed pattern is likely to be observed in the world.

We detect some cracks in the deductive orthodoxy, but findings-first papers will nonetheless face an uphill battle if judged by deductive criteria (Park, 2009). In some instances, the criteria are not far afield from those described earlier insofar as the research maps onto some a priori theoretical structure and can produce a belief shift about construct-to-construct links and the judged importance of those links. For example, Huber, Payne, and Puto’s finding of violations of regularity led to the rejection of broad classes of attraction models that could not accommodate the “attraction” effect. If a puzzle or effect is important enough or generative enough, reviewers should not require that the processes underlying the effect be pinned down in order for the research to be published.

In other empirical approaches we have described, findings and phenomena may have a less obvious a priori mapping to theory (Haig, 2005; Wegner, 1992). Although we strongly

endorse discovery-oriented research, questions of validity and robustness loom large. On the one hand, questionable individual data points may eventually aggregate to produce a robust and powerful meta-analytic conclusion. On the other hand, evidence indicates that many initial findings either fail to replicate or diminish in size over time (Hubbard & Armstrong, 1994; Ioannidis, 2005). Moreover, there is reason to question the self-correcting nature of science in those instances in which a paper reports a Spartan data point (Bertamini & Munafo, 2012; Ledgerwood & Sherman, 2012). Journals should not be an outlet for card tricks and one-off results.

For the author, our position walks a line between torment and trust. If a novel and important effect is uncovered, it should be allowed to proceed to publication without addressing reflexive calls to document moderation and mediation. Reviewers should also consider the logistical costs of replication, especially for field studies (see above). On the other hand, it is incumbent on the author to make a persuasive case that the effect is not only important but also real, with replication being the most natural route. However, even replication begs the question of scope. In our discipline, the broadest scope, and that which underlies empirical generation, is the use of multiple sets of data gathered across contexts and laboratories (Ehrenberg, 1995)—a criterion few would regard as reasonable for a single manuscript. It is not unreasonable for this criterion to take the form of within-laboratory replication, which minimally can be achieved via the use of separately analyzed replicates.

Aside from cost to the investigator, the degree to which replication should be enforced is a function of other subjective judgments that include: effect size, manipulation strength, realism, elegance, and plausibility. Reviewers are entitled to be skeptical when contrived manipulations

and convoluted designs produce small, head-scratching outcomes. A useful role for editors and AEs is in setting the stopping rule.

Non-deductive substantive contributions

In some research, concepts take a backstage role because the main focus is on documenting and measuring some particular substantive phenomenon. Although psychological or economic concepts are marshaled to interpret the dataset, the primary objective is to show, in a careful way, that a phenomenon exists, thereby paving the way for later studies of its explanation. Consumer researchers seem to scorn such papers as “atheoretical,” but such papers are alive and well among our colleagues in marketing science.

Lynch (2011a) noted that of the last 10 O’Dell Award winners, half fit this template, as exemplified by Mela, Gupta, and Lehmann’s (1997) demonstration that advertising decreases price sensitivity in the long run whereas price promotions have the opposite effect. Various authors and people in practice had argued that such effects might happen and, consequently, very little new was shown about the role of constructs. The main contribution is very carefully doing the econometrics to affect beliefs about links between *observables*. More generally, there is a rich tradition of “empirical generalizations” papers that carefully show some stylized facts about the marketplace (Ehrenberg 1995).

It is possible to add considerable confidence to beliefs about some property of the substantive system by convincing readers that the phenomenon in question has been precisely measured rather than introduce new concepts. For example, the rigor with which Mela et al.

conducted their investigation was critical in convincing readers of the validity of the result, allowing them to turn attention to the significance of the facts for marketing strategy.

Whereas Mela et al. examined some longstanding speculation on the part of practitioners and academics, other cases in this domain may spring solely from the authors' curiosity. Bronnenberg, Dubé, and Mela (2010) showed the surprising lack of effects of TiVo on consumer price elasticity and purchasing. In a three-year field study, some households were given a DVR and their shopping behavior was compared to those without one. The results showed that DVRs did not affect what people bought, contrary to the intuition that DVRs would make people skip commercials and become less price sensitive. Apparently, people with DVRs watch the vast bulk of shows live, and even when watching a recording they fail to engage the skip feature. One's judgment of the contribution of the study comes both from one's assessments of the care with which the phenomena were modeled, one's surprise at the findings, and one's judgment of how important DVRs are in the larger system of advertising and marketing communication.

Other disciplines seem to value non-deductive substantive contribution papers more than do consumer researchers. In economics, consider Woodward and Hall's (2012) analysis of consumer confusion in the mortgage market. They showed that brokers were exploiting borrowers' lack of knowledge and experience. They used empirical data on the compensation that borrowers pay to mortgage brokers for assistance from application to closing. They further showed that no rational model of search could explain the distribution of broker commissions, that borrowers sacrifice at least \$1,000 by shopping from too few brokers, and that borrowers who compensate their brokers with both cash and a commission from the lender pay twice as

much as similar borrowers who pay no cash, presumably because mixing cash and commission makes it more opaque to compare loans.

Such papers are equally important in consumer research but often struggle in the review process, despite possessing exceptional heuristic value in telling us what neglected topics in consumer research might prove to be truly important. For example, Dickson and Sawyer's (1990) now classic accounting of consumers' knowledge of supermarket prices battled to find a home in a premier journal — it was viewed as “merely descriptive.”

Criteria for reviewing. In evaluating non-deductive substantive research, belief shift may not always be the most important criterion. Rather, the reader makes a judgment about the degree to which the methods capture the phenomena of interest. Thus, in a scale development paper, evidence of reliability, convergent and discriminant validity, and nomological validity enter into the belief that the method “works.” Likewise, in an econometric paper, the reader needs to be convinced that the estimation method makes it possible to interpret the coefficients of interest. There remains a separate judgment of the importance of the phenomenon—which again is based on readers' judgments of the pervasiveness of the phenomenon and its psychological or economic significance. In the end, beliefs about relationships among observables in a real world substantive system are updated, not beliefs about construct to construct links or mappings from substantive phenomena to underlying constructs.

Concluding remarks

We have argued that consumer research has been dominated by the hypothetico-deductive approach, wherein researchers deduce from theory hypotheses about general construct-

to-construct links and then test those hypotheses with only limited concern for the substantive consumer behavior systems in which those links might be relevant. We respect the underlying logic of research in this vein and have published a fair amount of it ourselves. However, we also argue that a methodological orthodoxy has taken hold—a not unfair caricature of which includes mediation analysis even if meaningless (Zhao, Lynch, & Chen, 2010; Lynch, 2011b), a threshold number of experiments, a veneer of deductive logic, the reporting of higher-order interactions, and the inclusion of moderators. Authors, sensitive to r-quality, contribute to this rigidity when they conclude that journals will be unreceptive to departures from the dominant paradigm; reviewers validate authors' beliefs when they apply hypothetico-deductive criteria to manuscripts that employ “non-traditional” approaches.

In particular, we highlight benefits of other deductive approaches, particularly those for which the goal is to map theoretic constructs to substantive phenomena (e.g., memory principles can explain consideration sets) and those aiming to shift beliefs about observable-to-observable links (e.g., people given a raise who precommit save more for retirement because even if their budgets are tight they do not change their saving plan).

We also highlight the merits of approaches that trace a non-deductive path, where the chief insight or contribution comes from the findings rather than from the tightness of the explanation. Some such research aims to shift beliefs about whether a particular phenomenon exists (e.g., measuring intentions can actually cause people to buy expensive durables). Other such inductive and abductive research contributes by altering beliefs about construct-to-construct links (e.g., whether cues increase or decrease recall depends on category size). We also acknowledge that it would be a mistake to substitute one orthodoxy for another. Yet for these

papers to succeed and make it to publication, reviewers need to apply appropriate criteria in the review process. To that end, in this paper we have provided criteria for reviewing that we feel are most fitting to the particular research route taken. These criteria are also briefly summarized in Figure 2.

Insert Figure 2 about here.

We close with general remarks about the issue of how non-traditional manuscripts should be evaluated. At a general level, those who have lobbied for discovery and an emphasis on phenomena are remarkably consistent. Framed in terms of risk tolerance, Lehmann et al. (2011) and Wegner (1992) speak of the need to shift the focus from being right and avoiding error to not being wrong and pursuing truth. Cronbach (1986) argues that bias should favor the viability of an idea except when the costs of being wrong are intolerably high. Framed in terms of contribution, the emphasis is placed on insightfulness (Lehmann et al., 2011), “relevance, reality, and durability” (Rozin, 2001; see also Lehmann et al., 2011), and novelty, interestingness, and net progress (Rozin, 2009; see also Barwise, 1995).

Stated in terms of a currently popular research paradigm, Cacioppo, Semin, and Berntson (2004) compare the prevention focus of scientific realism, which skews toward deduction and emphasizes theoretical specification, critical tests, and rigor, to the promotion focus of scientific instrumentalism, which skews toward non-deductive approaches and emphasizes discovery and open-mindedness. Their overarching view aligns with our own. These alternative perspectives should be viewed as synergistic rather than antagonistic. In the end, precision is an objective but not at the expense of discovery. Rozin (2009) makes a similar point in the more specific context

of the review process. That is, the rules of “scientific hygiene” should not be enforced equally for the first versus 100th study on a particular topic. Doing so could prevent an idea from germinating.

As we have noted throughout this manuscript, there are two dilemmas that produce reviewer disagreement (with each other and with authors) about the contribution of manuscripts. The first is widely known but virtually irrepressible. Insofar as perceived novelty is a primary criterion, judgments of novelty will be subject to the robust, stubborn, and seductive plague of hindsight bias (Lynch 1998; Shavitt 2011; Slovic & Fischhoff, 1977), which prevents reviewers from perceiving any effect as surprising or, therefore, interesting. Although debiasing is not impossible, we are not sanguine about eliminating hindsight bias from the review process. It is the job of editors and associate editors to realize the reasons for reviewers having different priors. A paper can be important if it shifts beliefs of a significant group of readers even if others find it unsurprising.

The second dilemma pertains to risk tolerance and the need for journals to be “not wrong.” As Cichetti (1991) states, “with one review, editors “go with the flow,” with two, they “go with the low,” and with three, they “go with the mode” (see also Lynch, 1998; Park, 2010). Most often reviewers are lukewarm, in which case they are expected to ask for changes to the paper. Put another way, it is difficult for most reviewers to suggest changes to q-quality but easy to suggest changes to r-quality.

We close with an example of paper rumored to have been published by the editor over the objections of the majority of the review team. Goldstein et al. (2008) tested the effectiveness of signs put in hotel rooms that requested guests to participate in an environmental conservation

program and reuse their towels. They found that if the appeals on the signs highlighted that other guests participated in the program (e.g., “the majority of guests reuse their towels at least once during a stay”), hotel guests were themselves more likely to comply than when the traditional signs that just focused on the benefits of conservation were used. The authors concluded that descriptive norms are more effective than the standard industry message to induce compliance.

However, given the nature of the study, there was no way to verify that the effect was driven by or mediated by a focus on descriptive norms. Ideally one could examine which participants noticed the signs, assess how they rated them on normative and other aspects, and then determine how the normative ratings affected behavior. This, of course, was impractical because participants in this field study did not even know they were part of an experiment, getting the hotel to do what was needed to conduct this field study was likely already a Herculean feat, and trying to obtain follow up survey measures from hotel guests would have been extremely difficult, if not impossible. Yet, even without this additional information, the contribution of this study is significant, and was recognized with a *Journal of Consumer Research* Best Paper Award.

If we desire to increase q-quality, we need to move away from the flow-low-mode formula and move to a “champions” formula. We need at least one reviewer to be very enthusiastic about and champion a paper in order for q-quality to increase; r-quality is less likely to spark enthusiasm than is q-quality. This is especially important for phenomenon-driven research since judging whether a phenomenon is interesting or important to study inherently involves subjective evaluations and is likely to have lower inter-judge reliability. Editors should

strive to be guided not by the average of what reviewers recommend but by those who seem able to see the largest possible legitimate contribution of the work.

Figure 1: Four Categories of Inquiry

	Concepts First: Deduction	Findings First: Non-deduction
Intended Contribution to Conceptual Domain	Conceptual contributions via deduction	Conceptual contributions via non-deductive routes
Intended Contribution to Substantive Domain	Substantive contributions via deduction <ul style="list-style-type: none"> • Explaining substantive phenomena using theoretic constructs • Theory-based interventions to influence substantive systems. 	Non-deductive substantive contributions

Figure 2: Four sets of Criteria for Reviewing

	Concepts First: Deduction	Findings First: Non-deduction
Intended Contribution to Conceptual Domain	<p><i>Conceptual contributions via deduction</i></p> <ul style="list-style-type: none"> • Does this research shift your beliefs about construct-to-construct links? • Are these new beliefs important to theories of consumers even if they are not novel in other domains, e.g., do they overturn prior theoretical accounts of influential findings in consumer theory or have deep and broad explanatory power for consumer theory? • Do they have theoretical novelty. e.g., introduce a new moderator or mediator important for consumer theory? • Can the entire pattern of results (and not just a subset) be explained by a more compelling and equal or more parsimonious explanation? 	<p><i>Conceptual contributions via non-deductive routes</i></p> <ul style="list-style-type: none"> • Does this research shift your beliefs about construct-to-construct links? • Is there enough evidence that the effects are real or is further replication needed? (Consider here also the difficulty of studies done.) • Is the effect important enough that process is not necessary to pin down? • Are tests of moderation or mediation really needed?
Intended Contribution to Substantive Domain	<p><i>Substantive contributions via deduction</i></p> <p><u>Explaining substantive phenomena using theoretic constructs</u></p> <ul style="list-style-type: none"> • Does this research shift your beliefs about construct-to-phenomena links? • How important are those belief shifts? • How pervasive is the phenomena and what is its impact? (Are your prior beliefs or interests getting in the way here?) • Is the empirical demonstration clever and forceful? (Is it fair to ask the authors to change painstaking empirical work – does it have a fundamental flaw?) • Are the most essential features of the to-be-explained substantive phenomena captured? <p><u>Theory-based interventions to influence substantive systems</u></p> <ul style="list-style-type: none"> • Does this research shift your beliefs about observable-to-observable links – is it clear what operational treatments will produce effects? • Is there strong empirical evidence to support the internal validity of relationship between 	<p><i>Non-deductive substantive contributions</i></p> <ul style="list-style-type: none"> • Does this research shift your beliefs that a particular phenomenon exists? • Have the authors appropriately documented and measured the substantive phenomena, i.e., are the methods appropriate for capturing the phenomena? • Have we learned more about a substantive consumer topic? • Is the phenomenon important in the real world if true?

	interventions and real-world outcomes? Is the outcome clearly due to the intervention? <ul style="list-style-type: none">• Is the changed outcome important?• Is the effect size large enough to matter?	
--	---	--

References

- Aaker, D.A. & Keller, K.L. (1990). Consumer evaluations of brand extensions. *Journal of Marketing*, 54, 27-41.
- Aaker, J.L. & Lee, A. Y. (2001). 'I' seek pleasures and 'we' avoid pains: the role of self-regulatory goals in information processing and persuasion. *Journal of Consumer Research*, 28, 33-49.
- Abelson, R.P. (1995). *Statistics as principled argument*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Alba, J. W. & Chattopadhyay, A. (1985). Effects of context and part-category cues on recall of competing brands. *Journal of Marketing Research*, 22, 340-349.
- Ariely, D. & Levav J. (2000). Sequential choice in group settings: Taking the road less traveled and less enjoyed. *Journal of Consumer Research*, 27, 279-290.
- Barwise, P. (1995). Good empirical generalizations. *Marketing Science*, 14, G29-G35.
- Belk, R. W. (1984). Manifesto for consumer behavior of consumer behavior. *AMA Winter Educator Conference Proceedings*. 163-168.
- Bertamini, M. & Munafò, M. R. (2012). Bite-size science and its undesired effects. *Perspectives on Psychological Science*, 7, 67-71.
- Bower, G. H. (2007). Gordon H. Bower. In G. Lindzey & W. M. Runyan (Eds.). *A history of psychology in autobiography, Vol. 9*, (pp. 77-113). Washington DC: American Psychological Association.

- Brinberg, D. L., Lynch, J.G., Jr., & Sawyer, A.G. (1992). Hypothesized and confounded explanations in theory tests: A Bayesian analysis. *Journal of Consumer Research, 19*, 139-154.
- Brinberg, D.L. & McGrath, J. (1985). *Validity and the research process*, Beverly Hills CA: Sage Publications.
- Bronnenberg B., Dub J., & Mela C. (2010). Do digital video recorders influence sales? *Journal of Marketing Research, 47*, 998-1010.
- Cacioppo, J. T., Semin, G. R., & Berntson, G.G. (2004). Realism, instrumentalism, and scientific symbiosis. *American Psychologist, 59*, 214-223.
- Calder, B. J., Phillips, L.W., & Tybout, A.M. (1981). Designing research for application. *Journal of Consumer Research, 8*, 197-207.
- Cichetti, D. V. (1991). The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences, 14*, 119-186.
- Colquitt, J. A. & Zapato-Phelan, C. P. (2007). Trends in theory building and testing: A five-decade study of the academy of management journal, *Academy of Management Journal, 50*, 1281-1303.
- Cronbach, L. J. (1986). Social inquiry by and for earthlings. In D. W. Fisk & R. A. Shweder (Eds.), *Metatheory in social science* (pp. 83-107). Chicago: University of Chicago Press.
- Deighton, J. (2007). The territory of consumer research: Walking the fences. *Journal of Consumer Research, 34*, 279-282.
- Dickson, P. R., & Sawyer, A. G. (1990). The price knowledge and search of supermarket shoppers. *Journal of Marketing, 54*, 42-53.

- Dipboye, R.L. & Flanagan M.F. (1979). Research settings in industrial and organizational psychology: are findings in the field more generalizable than in the laboratory?" *American Psychologist*, 34, 141-150.
- Dunbar, K. (1999). The scientist In vivo: How scientists think and reason in the laboratory. In Magnani, L., Nersessian, N., & Thagard, P. (Eds.), *Model-based reasoning in scientific discovery*, (pp. 89-98), Plenum Press.
- Ehrenberg, A. S. C. (1995). Empirical generalizations, theory, and method. *Marketing Science*, 14, G20-G28.
- Ellison, G. (2002). Evolving standards for academic publishing: A q-r theory. *Journal of Political Economy*, 110, 994-1034.
- Folkes, V. S. (2002). Presidential address: Is consumer behavior different? In S. M. Broniarczyk & K. Nakamoto (Eds.), *Advances in consumer research*, Vol. 29 (pp. 1-4), Valdosta: Association for Consumer Research.
- Goldenstein, N.J., Cialdini, R.B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, 35, 472-482.
- Haig, B. D. (2005). An abductive theory of scientific method. *Psychological Methods*, 10, 371-388.
- Harford, T. (2011). *Adapt: Why success always starts with failure*. New York: Farrar, Straus, & Giroux.
- Hirschman E. C. (1986). Humanistic inquiry in marketing research: Philosophy, method, and criteria. *Journal of Marketing Research*, 23, 237-249.

Holbrook, M. B. (1985). Why business is bad for consumer research: The three bears revisited.

E. C. Hirschman & M. B. Holbrook (Eds.), *Advances in consumer research*, Vol. 12 (pp. 145-1), Provo: Association for Consumer Research.

Hubbard, R., & Armstrong, J. S. (1994). Replications and extensions in marketing: Rarely published but quite contrary. *International Journal of Research in Marketing*, 11, 233-248.

Huber, J., Payne, J.W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of Consumer Research*, 9, 90-98.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2, 696-701.

Jacoby, J. (1976). Consumer psychology: An octennium. *Annual Review of Psychology*, 27, 331-358.

Krishna, A. (2011). Can supporting a cause decrease donations and happiness? The cause marketing paradox. *Journal of Consumer Psychology*, 21, 338-345.

Ledgerwood, A., & Sherman, J. W. (2012). Short, sweet, and problematic? The rise of the short report in psychological science. *Perspectives on Psychological Science*, 7, 60-66.

LeDoux, J. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. Simon & Schuster: New York.

Lehmann, D. R. McAlister, L., & Staelin, R. (2011). Sophistication in research in marketing. *Journal of Marketing*, 75, 155-165.

- Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, 65, 272-292.
- Lynch, J.G., Jr. (1982). On the external validity of experiments in consumer research. *Journal of Consumer Research*, 9, 225-239.
- Lynch, J. G., Jr. (1998). Presidential address: Reviewing. In J.A. Alba & J. W. Hutchinson (Eds.). *Advances in consumer research*, Vol. 25 (pp. 1-6), Provo, UT: Association for Consumer Research.
- Lynch, J.G., Jr. (1999). Theory and external validity. *Journal of the Academy of Marketing Science*, 27, 367-376.
- Lynch, John G., Jr. (2011a), "Substantive consumer research," in *Advances in Consumer Research*, Vol. 38, ed. Darren W. Dahl, Gita V. Johar, and Stijn M.J. van Osselaer, Duluth, MN: Association for Consumer Research.
- Lynch, J.G., Jr. (2011b). Meaningless mediation. *Society for Judgment and Decision Making Newsletter*, 30 (3). <http://sjdm.org/newsletters/11-oct.pdf>
- Lynch, J.G., Jr., & Ariely, D. (2000). Wine online: Search costs affect competition on price, quality, and distribution. *Marketing Science*, 19, 83-103.
- Lynch, J.G., Jr. & Zauberman, G. (2006). When do you want it? Time, decisions, and public policy. *Journal of Public Policy and Marketing*, 25, 67-78.
- MacInnis, D. J. (2011). A framework for conceptual contributions in marketing. *Journal of Marketing*, 75, 136-154.

- MacInnis, D. J. & Folkes, V. S. (2010). The disciplinary status of consumer behavior: A sociology of science perspective on key controversies. *Journal of Consumer Research*, 36, 899-911.
- Madrian, B. C. & Shea, D.F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics*, 116, 1149-87.
- Maoz, E. & Tybout, A. (2002). The moderating role of involvement and differentiation in the evaluation of brand extensions. *Journal of Consumer Psychology*, 12, 119-131.
- McGrath, J. E. & Brinberg, D. (1983). External validity and the research process: A comment on the Calder/Lynch dialogue. *Journal of Consumer Research*, 10, 115-124.
- McGrath, R. G. (2011). Failing by design. *Harvard Business Review*, 89, 76-83.
- McNally, A., Prier, C. K., & MacMillan D. W. C. (2011). Discovery of an α -amino C-H arylation reaction using the strategy of accelerated serendipity. *Science*, 334, 1114-1117.
- Mela, C. F., Gupta, S., & Lehmann, D. R. (1997). The long-term impact of promotion and advertising on consumer and choice. *Journal of Marketing Research*, 34, 248-261.
- Morwitz, V.G., Johnson, E., & Schmittlein, D.C. (1993). Does measuring intent change behavior? *Journal of Consumer Research*, 20, 46-61.
- Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *Journal of Consumer Research*, 17, 263-276.
- Park, C. W. (2009). Goals for short articles. *Journal of Consumer Psychology*, 19, 570-573.
- Park, C. W. (2010). A two-step interactive review process. *Journal of Consumer Psychology*, 20, 105-106.

Park, C. W. (2012). Two types of attractive research: Cute research and Beautiful Research, *Journal of Consumer Psychology*, doi:10.1016/j.jcps.2012.005.0005

Raghubir, P. & Krishna, A. (1999). Vital dimensions: Antecedents and consequences of biases in volume perceptions. *Journal of Marketing Research*, 36, 313-326.

Reibstein, D.J., Day, G., & Wind, J. (2009). Guest editorial: Is marketing academia losing its way? *Journal of Marketing*, 73, 1-3.

Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Crown Business.

Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review*, 5, 2-14.

Rozin, P. (2009). What kind of empirical research should we publish, fund, and reward? *Perspectives on Psychological Science*, 4, 435-439.

Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Shavitt, S. (2011). Presidential address: What's new? Novelty in consumer research. In R. Ahluwalia, T. Chatrand, & R. Ratner (Eds.). *Advances in Consumer Research*, forthcoming.

Sheth, J. N. (1982). Consumer behavior: Surpluses and shortages. In A. Mitchell (Ed.). *Advances in Consumer Research*, Vol. 9 (pp. 13-16), Ann Arbor, MI: Association for Consumer Research.

- Shimp, T.A. (1994). Presidential address: Academic appalachia and the discipline of consumer research. In C.T. Allen and D.R. John (Eds.), *Advances in Consumer Research, Vol. 21*, (pp. 1-7), Provo: Association for Consumer Research.
- Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: Interplay of affect and cognition in consumer decision making. *Journal of Consumer Research, 26*, 278-282.
- Simonson, I., Carmon, Z., Dhar, R., Drolet, A. & Nowlis, S. (2001). Consumer research: In search of identity. *Annual Review of Psychology, 52*, 249-275.
- Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 544-551.
- Sternthal, B. (2010). ACR fellow's address: Three silos and a theory: Developing nomological networks to explain consumer behavior. In D. W. Dahl, G. V. Johar, and S. M. J. van Osselaer (Eds.). *Advances in Consumer Research, Vol 38*. Duluth, MN: Association for Consumer Research.
- Sternthal, B., Tybout, A. M., & Calder, B. J. (1987). Confirmatory versus comparative approaches to judging theory tests. *Journal of Consumer Research, 14*, 114-125.
- Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow (TM): Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*, 164-88.
- Wansink, B., & Van Ittersum, K. (2003). Bottom's up! The influence of elongation on pouring and consumption volume. *Journal of Consumer Research, 30*, 455-463.
- Wasserman, E. A., & Blumberg, M. S. (2010). Designing minds: How should we explain the origins of novel behaviors? *American Scientist, May-June*, 183-185.
- Watts, D. J. (2011). *Everything is obvious: Once you know the answer*. Crown Business.

- Wegner, D. M. (1992). You can't always think what you want: Problems in the suppression of unwanted thoughts. In M. Zanna (Ed.), *Advances in experimental social psychology*, (Vol. 25, pp. 193-225). San Diego: Academic Press.
- Wells, W. D. (1993). Discovery-oriented consumer research. *Journal of Consumer Research*, 19, 489-504.
- Woodward, S. E. & Hall, R. E. (2012), "Diagnosing consumer confusion and sub-optimal shopping effort: theory and mortgage-market evidence," *American Economic Review*, forthcoming.
- Yadav, M. S. (2010). The decline of conceptual articles and implications for knowledge development. *Journal of Marketing*, 74, 1-19.
- Yoon, Y., Gürhan-Canli, Z., & Schwarz, N. (2006). The effect of corporate social responsibility (CSR) activities on companies with bad reputations. *Journal of Consumer Psychology*, 16, 377-390.
- Zhao, X., Lynch, J., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37, 197-206.