A Critical Review of Question-Behavior Effect Research

UTPAL M. DHOLAKIA
Rice University

November 5, 2009

* Utpal M. Dholakia is the William S. Mackey and Verne F. Simons Distinguished Associate Professor of Management at the Jesse H. Jones Graduate School of Management, Rice University, 6100 Main Street, Houston, TX 77005-1892 (dholakia@rice.edu). I would like to thank Hanie Lee and Song-Oh Yoon for comments on earlier versions of this chapter.

A Critical Review of Question-Behavior Effect Research

Abstract

This chapter reviews research on the *question-behavior effect*, the phenomenon that asking questions influences respondents' behavior. Two distinct research streams, the *self-prophecy effect*, concerned with socially normative behaviors, and the *mere measurement effect*, dealing with purchase behaviors without socially normative significance, are identified. Despite the recent attempt at integration, it is argued that there are fundamental differences between the two effects. Distinctions are also drawn between lab-based and field-based mere measurement effects, and between normatively consistent and implicit attitude-driven, normatively inconsistent self-prophecy effects. Key studies, theoretical explanations, and moderators of each effect are discussed, potential unanswered questions and research opportunities are identified, and significant managerial and policy implications are highlighted.

Asking questions is the most common way of assessing an individual's internal states and predicting future behavior in social science research. In academic and applied settings, people are often asked to evaluate a particular object, issue, or organization, and to report their past and future behaviors. For instance, political parties, special interest groups, and media organizations poll potential voters regarding their positions on various issues and ask whether they will vote and for whom they will vote in an upcoming election. Marketing researchers ask consumers about their satisfaction with a particular product or a firm, their intentions to purchase products, and the degree to which they will recommend the product to others. Public health officials survey individuals about the frequency with which they perform various health-enhancing (e.g., wearing sunscreen, exercising, etc.) and risky (e.g., smoking, having unsafe sex, using drugs, etc.) behaviors. Economists are interested in eliciting the employment characteristics of citizens and their future outlooks toward consuming and saving.

In all of these cases, researchers implicitly assume that responding to questions will not have any subsequent influence on the individual. However, in testing its veracity over the last three decades or so, research has shown time and again that this assumption is tenuous at best, and invalid in many cases. Studies have found that answering questions, for example, through surveys, influences respondents in a variety of ways and through different psychological processes. The research area of the "Question-Behavior Effect" examines the short- and long-term psychological and behavioral effects of answering questions.

Virtually all Question-Behavior Effect (QBE) research can be traced to Sherman's (1980) study on the "self-erasing nature of errors of prediction." In a series of experiments, Sherman studied the ability of individuals to predict their future socially desirable actions. He found two

consistent results. First, when asked about a future socially normative behavior, study participants significantly over-predicted the degree to which they would perform it when compared to a control group that was simply given the opportunity to enact the behavior without being questioned. For instance, in one study, 47.8% of those asked to volunteer their time for a charitable cause predicted they would do so, but in reality only 4.2% of the control group volunteered. Second, respondents subsequently behaved in ways consistent with their over-predictions, that is, they acted in socially normative ways to a greater extent than the control group. In the charitable cause study, 31.1% of the respondents who were asked actually volunteered their time for the cause. Thus respondents' errors of behavioral prediction were self-erasing. Sherman's (1980) conclusion from the studies was: "When you look before you leap or predict behavior before you behave, the leaping and the behavior are likely to be altered; and indications are that the behavior will become more socially desirable and morally acceptable." (p. 220).

Since Sherman's paper, there have been dozens of studies examining the QBE. Researchers have replicated the effect in different settings, demonstrating its occurrence for socially desirable behaviors such as recycling, voting in elections, and donating to one's alma mater; socially undesirable behaviors such as gender stereotyping, using drugs, and skipping class; and neutral or "normatively ambiguous" (Spangenberg, Greenwald and Sprott 2008) behaviors, such as consumer purchases and relationships with firms. Studies have examined the magnitude and scope of the QBE; for example, its effect size (in the case of the self-prophecy effect), its occurrence for different types of questions, and its temporal pattern, i.e., how it evolves and how long it lasts. Researchers have also explored the underlying processes, discovering a number of boundary conditions, and considered its managerial, policy, and

consumer welfare implications (see Sprott et al. 2006, and Fitzsimons and Moore 2008, for recent reviews).

Objectives of this critical review

One noteworthy criticism of extant QBE research is that it is spread among different disciplines with oftentimes little information exchange between the areas. Researchers have used a number of different terms, studied different types of questions, and investigated a variety of behaviors. There are also variations in response modality, methodology, and theoretical explanations across the QBE studies. Table 1 summarizes these distinctions, showing the diversity of QBE research.

[Insert Table 1 about here]

Research on the mere measurement effect and the self-prophecy effect – the two dominant streams of QBE research – proceeded virtually independently for more than a decade before researchers in these areas acknowledged each other (Sprott et al. 2006). Recently, prominent researchers from the two camps have called to merge these research streams, proposing that:

> "We now find ourselves at a point where two once-independent groups of scholars have agreed to travel together towards an understanding of this phenomenon, as opposed to following separate, parallel paths… The beginning of this journey is to adopt formally a new descriptor for previously reported self-prophecy and mere-measurement effects. In particular, we encourage the use of the label *question-behavior effect*…By acknowledging similarities in the literature and adopting a shared, single label for related, observed effects, we can step back and take a comprehensive look at the broader set of phenomena we have observed and the proposed explanations for these phenomena." (Sprott et al. 2006, p. 129).

This call for integration is admirable, and there is no question that both mere measurement and self-prophecy researchers are studying related issues. However, it is still not clear how far to take the integration of these research streams. Should the terms "mere

measurement effect" and "self-prophecy effect" be abandoned entirely? In my view, the answer depends on the degree of overlap between these effects with respect to when, how, and perhaps most importantly, *why* they occur.

The current state of QBE research is that it is disjoint; there is sore need not only for understanding degrees of similarities and differences between the mere measurement effect and the self-prophecy effect, but also for a consolidation of what is already known. Relatedly, a consideration of its implications is needed along with an elaboration of the gaps in our understanding and the promising next steps to advance our state of knowledge regarding question-behavior effects. The current chapter seeks to accomplish these objectives.

Its purpose is to provide a critical review of the QBE research area. In the next two sections, I first give attention to research on the mere measurement effect and next to research concerning the self-prophecy effect. Despite the recent integration attempt, I argue that there are fundamental differences between the two effects. To prevent conceptual confusion and stimulate knowledge development, researchers are advised to specify which effect they are studying and position their contributions and findings to the relevant effect when designing research studies and interpreting their findings. For example, mere measurement studies concern purchase behaviors which are normatively neutral in the sense that acting or not acting does not have socially desirable or undesirable elements from the consumer's standpoint. In contrast, self-prophecy studies exclusively examine socially normative behaviors. Consequently, the primary explanation provided for one phenomenon, say, the self-prophecy effect, is of limited utility in explaining why the mere measurement effect occurs, and vice versa.

I also draw and elaborate on the distinctions between lab-based and field-based mere measurement effects, and between traditional self-prophecy research and the nascent, evolving,

and potentially important research on effects of surveys on risky behaviors of adolescents (Fitzsimons and Moore 2008). Within each research stream, I review and discuss key published studies, examine different theoretical explanations, and describe known moderators. Table 2 summarizes the key distinctions between the four types of QBE research discussed here.

[Insert Tables 2 and 3 about here]

Throughout the chapter, I highlight potential unanswered questions and research opportunities to advance our knowledge of the QBE, which are summarized in Table 3. Finally, I discuss the significant managerial and policy implications of QBE research. I argue that the findings and its potential consequences merit widespread attention and thought among researchers and practitioners who use a questioning methodology to elicit information from individuals for any purpose.

THE MERE MEASUREMENT EFFECT

Introduction to the mere measurement effect research stream

This line of research, conducted primarily by marketing and consumer researchers, has focused on effects of questioning on consumers' purchase behavior. The original study was that of Morwitz, Johnson, and Schmittlein (1993) who coined the term "mere measurement effect" (MME) based on the observation that *merely measuring* purchase intent of consumers impacts their purchasing behavior. Using panel survey datasets, they studied purchases of automobiles and personal computers among panelists who had completed a purchase intentions survey and those who did not participate during the six months afterward. They found significant and practically meaningful increases in both cases. For PCs, asking intent increased the purchase rate by 18% (3.80 % for non-respondents vs. 4.48% for respondents), and for automobiles, the

increase was 37% (2.4% vs. 3.3%).

Review of mere measurement effect research studies

Building on this initial study, Fitzsimons and Morwitz (1996) shifted the frame of analysis from product category to brand. They argued that asking a purchase intentions question about a product category (without referring to individual brands) activates the category in proportion to the prior accessibility of brand cognitions. Thus, a brand that is more accessible previously is more likely to be activated in memory and influence behavior. Results of their study showed that among current car owners, increase in choice incidence accrued to their current car brand. For example, Saab owners became more likely to purchase another Saab. In contrast, those who did not own a car were more likely to purchase a popular brand with a large market share.

Since then, a number of studies have demonstrated this effect. Morwitz and Fitzsimons (2004) showed its occurrence in participants' choice of Canadian candy bars through laboratory experiments. In one of their studies, for example, the choice share of a target candy bar increased from 29.7% when purchase intentions were not measured to 54.1% when purchase intentions were measured. In another paper, students who were asked (vs. not asked) their likelihood of flossing teeth reported greater instances of teeth flossing in a subsequent two-week period (Levav and Fitzsimons 2006; see also Williams, Fitzsimons and Block 2004). Janiszewski and Chandon (2007) demonstrated the MME's occurrence through lab-based studies that involved choosing ice-cream treats and candy bars.

Dholakia and Morwitz (2002) studied the effects of measuring customer satisfaction. In a field experiment conducted by a large US financial services firm, one customer group of 945 participated in a telephone-based satisfaction survey regarding the firm and its products, whereas

a comparable group of 1,064 was the control. Both groups were withheld from the firm's direct marketing activities for a year after the survey, and their behaviors and profitability were tracked during this time. Results showed that survey participants owned significantly more accounts (5.45 vs. 3.39), had a defection rate that was less than half (6.6% vs. 16.4%), and were significantly more profitable ($107.8 per month vs. $97.2 per month) than the control group. These differences were persistent. Survey participants continued to open new accounts at a faster rate and to defect at a much slower rate than non-participants, even a year afterward. As I argue in detail later, this and other field-based, long-term mere measurement studies are different in several crucial aspects from lab-based studies.

In cooperation with the leading French web-based grocer, Chandon, Morwitz and Reinartz (2004) studied the incidence, timing, and profitability of online grocery purchases made by consumers whose purchase intentions were measured and those of a control group. They found that measuring intentions led to an increased likelihood of repeat purchase and a shortened length of time before the first repeat purchase. For instance, one month after the survey, 9% of the control group and 20% of the surveyed group had made at least one repeat purchase from the site. However, both effects decayed rapidly after three months. Nevertheless, they found persistent gains in customer profitability because the accelerated purchases of the first three months led to faster subsequent purchases in the remainder of the nine-month period of their study.

Dholakia, Morwitz and Westbrook (2004) explicitly compared those who expressed *medium* and *low* levels of satisfaction in the firm's survey, in addition to those who expressed *high* levels of satisfaction. Their results revealed that in comparison to a control group, all three surveyed groups exhibited *more* purchase and relational behaviors. In their study of customers of

a large US automotive services provider, Dholakia, Singh, and Westbrook (2009) found that survey participants *delayed* their very next visit to the firm's stores, even when expressing high satisfaction, but accelerated later service visits. Through a lab experiment, they explained these results through increased service comprehensiveness perceptions among survey participants.

Research has also shown that asking hypothetical questions (e.g., "If strong evidence emerges from scientific studies that cakes, pastries, etc. are not nearly as bad for your health as they have been portrayed to be, and may have some major health benefits, what would happen to your consumption of these items?") can have a significant biasing effect on behavior (Fitzsimons and Shiv 2001). The percentage of respondents who chose cake over fruit increased significantly if respondents had been asked a hypothetical question about the benefits of baked goods an hour earlier. The authors proposed that such questions enhance accessibility of cognitions related to the false proposition(s) provided in the hypothetical question, leading the decision maker to behave in ways that are consistent with these activated cognitions. Consistent with this process, an increase in cognitive elaboration increased the contaminative effects of hypothetical questions, especially when the hypothetical questions were perceived as relevant to the decision (Fitzsimons and Shiv 2001). However, when confronted with the possibility that the hypothetical question may have guided behavior, they denied this association, suggesting the operation of an automatic process.

Studies have also examined the effects of informing consumers prior to a service encounter that they will be asked to evaluate it afterwards. Across a number of studies, results showed that forewarning customers leads them to provide less favorable quality and satisfaction evaluations and reduces their willingness to purchase and recommend the evaluated service (Ofir and Simonson 2001; Ofir, Simonson and Yoon 2009). Demonstrating its robustness, this effect

occurred in cases where actual service quality was either low or high, and even after participants were explicitly instructed to consider positive and negative aspects of the service (Ofir and Simonson 2001). Its impact on the consumer's evaluation of the service was enduring, lasting several days after the service encounter (Ofir et al. 2009). The authors concluded that "expecting to evaluate the store's service appears to change the actual shopping experience and promote a more thorough evaluation process" (Ofir et al. 2009, p. 14).

Anticipating answering questions before actually answering them and simply answering questions without forewarning have dramatically different effects on behavior that should be examined further in future research. For example, it would be interesting to contrast the behavior of customers who anticipate and later complete satisfaction surveys, customers who only complete satisfaction surveys (without prior anticipation), and a control group, within a single study, to determine the relative magnitudes of the different effects on behavior. Next, differences between lab-based and field-based mere measurement research are considered.

Lab-based vs. field based mere measurement research

The reviewed studies indicate that researchers have used starkly different approaches and measures to study the MME. Two distinct methodological approaches can be discerned in these studies, which I refer to as "lab-based" and "field-based" mere measurement effects in this chapter. I argue that this difference has significant bearing on how the MME unfolds with respect to its scope and persistence, and the process(es) driving the effect.

Lab-based mere measurement research is typically conducted via controlled experiments in the laboratory within relatively short time frames. The gap between question and behavior is of the order of minutes or hours. These studies often utilize novel stimuli such as Canadian candy bars (Morwitz and Fitzsimons 2004) and European ice cream treats (Janiszewski

and Chandon 2007) to be able to manipulate the subject's attitudes and to prevent contamination from prior knowledge. The stimuli are low-priced, frequently purchased food items. The dependent measure usually studied in this research is *choice* (selection of one option from a set of alternatives) or *purchase likelihood* ("How likely are you to purchase [the product]?") in most cases. Less often, actual purchase behavior of participants is measured (e.g., Janiszewski and Chandon 2007, Experiment 4), but even in these cases, the amounts involved are a few dollars.

Field-based mere measurement studies are different in all these respects. They are conducted in the noisy real-world environment with existing customers of a firm over longer periods of time. The gap between question and behavior is of the order of weeks, months, and even years (e.g., Borle et al. 2007; Chandon et al. 2004; Dholakia and Morwitz 2002; Morwitz et al. 1993). Unlike lab study subjects who are unacquainted with the product beforehand, participants not only know the firm and its products well, but they have an ongoing relationship with it. Questioning therefore has the potential to influence a wider range of behavior, including repurchasing, complaining, communicating, word-of-mouth, and identity-expressive behaviors. Because of an ongoing relationship, effects of questioning have the potential to occur over an extended period of time, i.e., weeks or months post-survey.

Field-based studies have covered a number of industries, from automobiles and PCs (Morwitz et al. 1993), to financial services (Dholakia and Morwitz 2002), online grocery retail (Chandon et al. 2004), and automotive maintenance services (Borle et al. 2007; Dholakia et al. 2004). Thus, the contexts tend to involve products with a wider range of prices and levels of consumer involvement. Additionally, field-based studies measure actual behaviors and performance metrics such as customers' defection rates, profitability, frequency of repeat purchase, and number of items purchased per visit. A broader range of customer behaviors is

assessed in these studies relative to lab-based studies. Many of these behaviors are effortful or difficult to implement for consumers, making occurrence of the question-behavior effect in a sustained and broad-based manner even more noteworthy (Gollwitzer and Oettingen 2008).

In the current chapter, I propose that because of these stark differences, the MME phenomenon studied in the laboratory is *essentially different* from the effects observed in the field. Most existing research has tended to ignore or downplay the distinctions between these effects, implicitly assuming that findings in one setting and the reasons for their occurrence fully extrapolate to the other setting. However, I propose that the theoretical explanations for why the MME occurs in the field, and therefore its boundary conditions, the means of attenuation, and practical implications, are, by and large, different from the lab-based MME.

Theoretical explanations for the mere measurement effect

In the decade and a half since the Morwitz, Johnson, and Schmittlein (1993) article, considerable scholarly effort has gone into explaining why the MME occurs. In this section, I discuss theoretical explanations for the lab-based MME first, followed by what is known about why the field-based MME occurs.

Explanations for occurrence of the lab-based mere measurement effect

Questioning increases accessibility of attitudes. The most widely cited explanation for the MME is based on increased accessibility of attitudes because of responding to questions. Researchers have generally relied on *self-generated validity theory* (Feldman and Lynch 1988; Simmons, Bickart and Lynch 1993) as the basis for explaining why the effect occurs.

According to this theoretical explanation, when survey respondents are asked a question – e.g., their satisfaction evaluation – many are unlikely to have formed such a judgment spontaneously beforehand or even given much thought to the issues pertaining to such a question

(Weiner 1985). Upon being asked, the respondent engages in thoughtful processing, constructs a response, and provides it to the questioner. The cognitive processing and articulation of the judgment increase its subsequent accessibility, resulting in behaviors that are consistent with the expressed judgment (Alba, Hutchinson and Lynch 1991; Kardes, Allen and Pontes 1993; however see Converse 1970, for a different perspective). During the time that accessibility of responses remains greater, it affects post-survey behavior (e.g., Fitzsimons and Williams 2000; Morwitz and Fitzsimons 2004).

This explanation is supported by a number of lab studies. As an example, in a series of experiments, Morwitz and Fitzsimons (2004) provided process-based evidence that responding to a purchase intention question increases accessibility of the attitude towards the behavior. Participants were asked to form attitudes about Canadian candy bars (that they were unfamiliar with), then list reasons for purchasing or not purchasing a particular bar, and finally indicate whether they would purchase *any* bar (general purchase intention). Participants who provided their general purchase intention were more likely to choose a particular bar if they had listed positive reasons for purchasing it, and vice versa. They were also more likely to recall the more accessible brand and could judge it as good with much more speed/using less time. Across the studies, respondents were more likely to choose options toward which they held positive and accessible attitudes, and less likely to choose options with negative accessible attitudes.

There is also evidence that the increased accessibility of attitudes upon questioning occurs largely through an *automatic process* (Fazio et al. 1986). Fitzsimons and Williams (2000) tested the extent to which the MME occurs because consumers carefully consider making a purchase among various brands in the category or because the question automatically invokes category members, heightening their pre-existing accessibilities. In lab studies using a process-

dissociation procedure to separately estimate contributions of automatic and effortful processing[1], the authors demonstrated that the change in respondents' behavior was more than three times as much due to automatic activation of the cognitive structure in which that information is contained as the effect of effortful processing. Note that such a process requires the existence of a well-learned set of cognitions regarding category members (Fazio et al. 1986) implying that product category experience may strengthen the relative role of automatic processing.

Interestingly, some of the field-based studies have also relied on increased attitude accessibility to explain their results. Dholakia and Morwitz (2002) and Chandon et al. (2004) both argued that accessibility of their respective measures explained the effects that they observed. Dholakia, Morwitz and Westbrook (2004) also found consistent results among long-standing customers in their field study. When expressing dissatisfaction in the survey, such customers were likely to behave less relationally when compared to control customers, indicating their behavior was in line with their now more accessible negative attitudes.

Despite these findings, none of the field-based studies have provided process-based evidence for operation of increased attitude accessibility. In the social psychology literature, increased accessibility of information is viewed as an automatic and short-term phenomenon (e.g., Bargh et al 2001; Gilbert and Hixon 1991; Wyer and Srull 1989), generally lasting for a few minutes. The extent to which it also influences long-term consumer behavior after questioning requires more careful validation (see also Levav and Fitzsimons 2006, and Sprott et al. 2006, for similar views). One way to do so would be to adopt Morwitz and Fitzsimons'

[1] The process dissociation procedure (Jacoby 1991) involves task performance under two conditions: an "inclusion" condition in which the automatic and effortful processes work synergistically to contribute to performance, and an "exclusion" condition in which they oppose each other. The difference in performance between these conditions provides an estimate of the contribution of each process.

(2004) experimental paradigm, executing it in the field over a longer period of time.

Additionally, some findings from lab-based studies done by Chapman (2001) are inconsistent with increased attitude accessibility. First, he showed that measurement of purchase intentions influenced behavior toward novel products (see also Janiszewski and Chandon 2007). In these cases, increased attitude accessibility is inapplicable because it is impossible for consumers to have attitudes when they have not encountered the product before. Second, Chapman (2001) found that when consumers repeatedly respond to an intent question regarding a novel product, they exhibit an increase in attitude toward the product, but this is not accompanied by an increase in positive thoughts regarding it, again ruling out an attitude-driven explanation. Finally, he found a stronger MME for purchase intentions than for product attitudes. Assuming that answering questions about attitudes directly increases their accessibility to a greater extent than answering intent questions without attitude elicitation, this pattern of findings is anomalous.

To summarize, increased attitude accessibility is currently the leading explanation for the lab-based MME. Although some field-based studies have invoked it, compelling evidence to support the thesis that attitudes made accessible through questioning remain available for weeks, months, or years afterward is still lacking. Additionally, evidence from some lab-based studies indicates that other forces are may be at work in conjunction with, or in place of, increased attitude accessibility when individuals are questioned.

Questioning results in behavior simulation and increases response fluency. Another explanation which can account for the MME's occurrence proposes that when responding to survey questions, individuals utilize *simulation*, imagining the target behavior and the specific circumstances under which they might engage in it. When the time comes for enactment, the

simulated action sequence is easier to retrieve than alternative unimagined actions, and it directs the respondent's behavior. Behavior simulation does not necessarily have to be elaborated on by the respondent; it can also occur automatically through unconscious processing such as by greater ease of fluency (Levav and Fitzsimons 2006). Sherman (1980) first proposed (but did not test) this explanation by suggesting that when questioned, respondents generate a *script*, or a mental representation of stereotyped event sequences for a particular situation (Abelson 1981). They then simply invoke the script when the situation actually occurs. Behavior simulation can explain the MME for novel behaviors, and a stronger MME for intentions than attitudes, which increased attitude accessibility is not able to explain (Chapman 2001).

As I discuss later in this chapter, behavior simulation is also conducive to explaining self-prophecy effects for socially normative behaviors in some cases (Spangenberg and Greenwald 1999). Finally, a related possibility is that responding to questions under certain circumstances (e.g., for short-fuse behaviors which have a limited window of opportunity for enactment) facilitates the production of an *implementation intention* specifying when, where, how and how long the actions will be carried out (Gollwitzer 1999). Prior research provides ample evidence that implementation intentions make it more likely that short-fuse behavior such as using a coupon before expiration will be enacted during the open window of opportunity (e.g., Dholakia and Bagozzi 2004). QBE research still has not explicitly considered the role of implementation intention formation upon questioning and its influence on behavior (however, see Goldstein et al. 2008, for comparison of effects of measuring behavioral intentions *or* forming implementation intentions on voter turnout). It could be that certain types of questions such as specific purchase intentions are more conducive to implementation intention formation because the respondent "fills in the blanks" regarding the contingencies under which the behavior will be enacted.

Levav and Fitzsimons (2006) provided empirical support for increased response fluency by testing boundary conditions for the MME's occurrence. In their *ease-of-representation hypothesis*, the authors posited that the effect of questioning on behavior is an increasing function of the ease with which the behavior is mentally represented. They argued that this is because questions about intentions lead to two related mental operations: representation of the target behavior and an assessment of how easily the representation came about. In their experiments, different manipulations that were designed to increase ease of mentally representing or simulating the behavior (described in detail in the section on moderators of the MME) increased the strength of measurement.

Janiszewski and Chandon (2007) tested the role of response fluency more directly, offering an explanation based on *transfer-appropriate processing*. They hypothesized that the redundancy in cognitive processes used to generate responses during the initial questioning and the cognitive processes used to decide whether to engage in the behavior at a later time creates a processing fluency favoring actions consistent with the original response. Through a carefully executed series of eight lab studies, they provided evidence that processing fluency contributed to the MME beyond either attitude or information accessibility.

Nevertheless, at least three limitations of the studies reported by Levav and Fitzsimons (2006) and Janiszewski and Chandon (2007) are worth pointing out. First, most studies have elicited behavioral expectations as the ultimate dependent variable (e.g., purchase likelihood) instead of actual behavior. Thus, it is not clear whether the pattern of results found in the studies supporting the process would hold for actual behaviors (see Sherman 2008, for an elaboration of this issue). Second, the studies were limited to simple behaviors (e.g., flossing, Levav and Fitzsimons 2006) and the purchase of low-priced products (ice cream treats and candy bars,

Janiszewski and Chandon 2007). Consequently, it remains unknown whether processing fluency can produce a MME for complex behaviors involving self-regulation or for expensive products. Finally, although Janiszewski and Chandon (2007) cited evidence from cognitive psychology suggesting that processing-fluency effects can persist for weeks or even months after the initial exposure, direct evidence for the longer-term existence of processing fluency effects is lacking. Note, however, that Levav and Fitzsimon's (2006) studies did span weeks and examined processes indirectly through consideration of boundary conditions. In summary, as it presently stands, behavior simulation/increased response fluency has garnered considerable evidence for lab-based MMEs, but little evidence for the field-based MME.

***An ideomotor process explanation*.** Related to the behavioral simulation account, Spangenberg, Greenwald and Sprott (2008) recently proposed that the *ideomotor effect*, defined as the perceptual image or idea of action that facilitates initiation of that action when no other contradictory idea is present in the mind (James 1950), can help explain both mere measurement and self-prophecy effects. They suggested that a question activates a perceptual image or idea of the action. The activated image guides future performance of behavior. According to the authors, an ideomotor perspective may account for question-behavior effects for a variety of behaviors since many everyday actions are likely to have clear ideomotor representations. This explanation currently remains empirically untested. Assuming its occurrence, its scope and extent of occurrence are also unknown, offering promising future research possibilities.

Attitude polarization. Morwitz and Fitzsimons (2004) also proposed an alternative explanation for the MME limited to conditions when intentions are repeatedly assessed. They suggested that asking an individual their behavioral intentions could activate the node for the general behavior and thus access his or her attitude. To the extent that accessing the node for the

general behavior functions in the same manner as repeated expression of a response, they hypothesized there should be a *polarizing effect* on initial attitudes for highly accessible choice options and a corresponding change in choice favoring these options. Specifically, those who express high initial levels of intent should have higher repurchase rates and those with low initial levels of intent should have lower purchase rates the more often their intent is measured relative to a control group (Morwitz et al. 1993).

Morwitz, Johnson and Schmittlein (1993) found some evidence supporting this explanation, but only for those respondents who reported low initial intent levels in their studies. They did not find this effect for those expressing high initial intent levels. Through lab experiments, Morwitz and Fitzsimons (2004) were able to rule out attitude polarization as an explanation for their findings. However, Dholakia et al. (2009) did find higher overall satisfaction levels among consumers who had been surveyed and had reported their satisfaction with a particular service visit either four or nine months later, indicative of polarization. To date, this is not a favored explanation for the MME. However, it may explain the effects of repeated questioning on some respondents, and could have practical implications, such as providing guidelines to firms regarding how frequently they should survey their customers. For a satisfied customer base, for example, structured questioning every one or two years may be beneficial. More studies are needed to investigate this explanation in depth.

Explanations for occurrence of field-based mere measurement effect

Although a number of field-based studies have invoked increased attitude accessibility, at least two other explanations are as, if not more, relevant for field-based MMEs.

Questioning generates positive inferences. One explanation consistent with self-generated validity theory (Feldman and Lynch 1988) is that participation in a survey conveys

favorable information about the firm. Survey participation leads customers to formulate positive inferences for example, “The firm values my opinions,” “It is making a bona fide effort to please me,” “It is caring and concerned about its customers in general,” etc., which influences the individual’s behaviors. This “positivity effect” (Dholakia et al. 2004) is applicable to marketing research surveys where the firm first identifies itself as the survey’s sponsor.

This explanation is consistent with consumer psychology theorizing that positive inferences formed on the basis of a single employee’s actions can influence the customer’s evaluation of the entire firm (Folkes and Patrick 2003). It is also congruent with the theory of *selective hypothesis testing* (Sanbonmatsu et al. 1998) which posits that individuals often form focal hypotheses regarding firms (e.g., the firm is a good organization to conduct business with) based on initially encountered evidence and using few information sources (or even a single source), which are used to guide interpretation of the gathered evidence and an assessment of the evidence’s validity (see also Ofir and Simonson 2001).

Dholakia, Morwitz, and Westbrook (2004) found that participation in a satisfaction survey conducted by an automotive services retail chain led customers to engage in service visits with greater frequency, purchase more services, and become more likely to redeem coupons, *even when they expressed dissatisfaction in the survey* when compared to a control group. Such a pattern of findings cannot be explained by increased attitude accessibility, which predicts reduced visits for the dissatisfied group (compared to the control). In explaining why adolescents perform risky behaviors such as using drugs after being questioned, Gollwitzer and Oettingen (2008) suggested that questioning may not only affect the implicit attitude’s state of activation but also increase its level of positivity, a notion that is consistent with the positivity effect.

However, unlike implicit attitudes, the positivity effect relies on the consumer’s

conscious and thoughtful cognitive processing regarding the firm's motives for conducting the survey. Through this process, survey participation enduringly changes the respondent's opinion regarding the firm. Relative to increased attitude accessibility, the positive inference account is better able to explain why firm-sponsored survey participation has a broad-based and long-lasting positive impact on customer behaviors that can be observed for weeks or months after the survey. However, it is not able to explain why the effect also occurs for unsponsored surveys such as those involving consumer panels (e.g., Morwitz et al. 1993) or those conducted by third-party organizations. Although plausible, to date, lab-based MME studies have not directly examined whether the positivity effect can occur in the lab setting, and under what circumstances.

Likewise, much remains to be known regarding the boundary conditions for positive inferences generated by customers even in the field. For example, it could be that when the firm conducts a survey by telephone (as in Dholakia et al's [2004] study), the respondent generates positive inferences because of the higher perceived cost and effort expended by the firm in questioning; in contrast, when the survey is done online, it may be perceived as less costly and effortful, and might not have a similarly positive effect on customer behavior (e.g., Johnson and Folkes 2007). Other factors such as the use of a professional interviewer, the opportunity to provide open-ended feedback and suggestions, and a follow up conversation explaining how the feedback given in the survey was acted on could each facilitate generation of positive inferences, enhancing positive behavioral effects from mere measurement.

Further understanding the specific cognitive process through which the positivity effect occurs is also a promising research opportunity. At least two possibilities seem evident. First, the positivity effect could occur because of an improved customer relationship with the firm. When

asked to participate in a satisfaction survey, customers may infer that such a request is a *relationship-enhancing attempt by the firm directed specifically at them*. In response to such an overture, customers may experience positive affect toward the firm, and reciprocate, leading to more favorable behavior than if they had not been asked to participate.

An alternative possibility is an *expectancy explanation*, which holds that customers selected for the survey attribute the firm's request to its customer orientation and its commitment to providing superior service. Since these are nearly universally-held beliefs about desirable characteristics of firms, customers may come to view the firm conducting the survey as a superior one seeking above its competitors to please and be responsive to its customers. As a result of their enhanced perceptions and expectations, these customers would display correspondingly stronger relational behaviors toward the firm.

Questioning generates a broad range of inferences. Rather than just positive inferences, a recent study by Dholakia, Singh and Westbrook (2009) proposed that answering firm-sponsored satisfaction surveys can produce inferences among respondents regarding issues beyond those specifically inquired about in the survey. They argued that because a firm's customers have a broad base of knowledge and a high degree of interest in the questioning, they possess both the motivation and the ability to generate a wide range of positive and negative inferences. Taking the perspective of the survey respondent as a thoughtful individual (Bradburn, Rips and Shevell 1989), such an explanation is consistent with recent social psychological research showing that individuals can spontaneously infer goals from verbal stimuli such as goal-implying sentences (Hassin, Aarts and Ferguson 2005), which can then affect behavior (Aarts, Gollwitzer and Hassin 2004).

To test this proposition, the authors examined the role of service comprehensiveness

inferences regarding an automotive quick lube service, which they defined as the customer's belief that more elements of the service were performed in addition to the oil change, resulting in a more thorough checkup of the vehicle. In a lab study, they found that satisfaction survey participants recalled more specific service elements, and reported receiving more complete service from the firm than non-participants. They also described a longitudinal field study conducted in cooperation with a national automotive services chain which showed that post-survey, consistent with inferences of service completeness, participants delayed their very next visit even when they reported being highly satisfied with the last one, but accelerated later service visits when compared to non-participants. This was the first MME study to show *contrasting valence* for satisfied customers, i.e., satisfied customers showed a negative behavior (i.e., delay in repurchasing the service) after being questioned.

Ofir and Simonson's (2001) research on the effects of expecting to evaluate on satisfaction evaluations provides additional evidence for inference-making by survey respondents. In understanding why expecting to evaluate led to more negative evaluations, they found support for a "negativity enhancement" explanation, whereby expecting to evaluate prior to questioning reinforced the consumers' tendency to focus on and overweigh the negative aspects encountered during the service. They argued that one reason for negativity enhancement is that consumers inferred that the service provider wanted them to offer negative and constructive criticisms, which is why they tended to focus on weaker or underperforming aspects of the service. Furthermore, they sought out negative aspects because such elements appeared to be much more diagnostic and therefore useful to the service provider. Thus, current findings indicate that forewarning customers they will be asked to evaluate satisfaction induces negative inferences (e.g., "Firms need improvement and I should provide constructive feedback") and

actual participation in a satisfaction survey can lead to positive inferences (e.g., "Firms care about customers"). Providing an over-arching theoretical explanation for these findings is an interesting future research avenue.

Relatedly, Lusk, McLaughlin and Jaeger (2007) demonstrated that survey respondents act in their own self-interest, responding to purchase intentions questions strategically by making inferences about how their responses will influence the product's future price, and the marketer's decision of whether to offer the product. Finally, the proposed process underlying Levav and Fitzsimons' (2006) "ease of representation" hypothesis also supports the possibility of inference-making by respondents. They suggested that when providing responses, respondents infer likelihood of the behavior's enactment from the ease with which they can represent it, which in turn spurs the formation of an implementation intention for the future.

The notion that consumers form a broad range of inferences due to survey participation, which could potentially change how they view the firm in a significant and durable fashion, is a drastically different explanation for the MME than increased attitude accessibility discussed earlier, which focuses on shorter-term activation of responses. The stark contrast between these explanations captures the fundamental conceptual differences between the lab-based and field-based approaches to studying mere measurement. Nevertheless, although the notion of inference-making is intuitively appealing, and receives support from social psychological theorizing and the described studies, it does raise the question of tractability. Potentially, survey respondents could simultaneously generate a multitude of inferences, some of which work in tandem, others that counteract, and still others that are unrelated to one another, in influencing behavior. Considerably more theoretical and empirical work is needed before we know the different types of inferences produced by surveys in the field and under what circumstances they occur.

Summary As the discussion in this section indicates, existing evidence strongly suggests that different processes or more likely combinations of processes may be at work when the mere measurement effect occurs in the lab than when it occurs in the field. Of the existing studies, a large proportion has focused on increased attitude accessibility, and this explanation appears amenable to playing a significant role in occurrence of lab-based, and potentially some role in the field-based, MME. Relatively few studies have focused on other explanations and in most cases, there are a handful of studies examining any one particular process. There is a great need for future studies not only to confirm occurrence of these other processes but more importantly to gain a deeper understanding regarding the conditions under which each process (or combination thereof) contributes to the MME's occurrence.

Moderators of the mere measurement effect

In this section, I review what is known about current boundary conditions for the MME. Most studies identifying boundary conditions have done so for field-based MMEs. These studies provide new and useful insights into the subtleties, scope, and processes underlying the MME. Very few studies have examined boundary conditions for the lab-based MME (Levav and Fitzsimons 2006 is a notable exception).

The consumer's experience with the product category. The consumer's product experience is a well-established moderator in MME research. Studies generally show that the MME diminishes with product experience. Morwitz et al. (1993) found this to be the case in their study. In the PC category, for example, measuring intent increased sales of PCs by 20.7% for experienced consumers, but by 45.3% for those without prior product experience. Similarly, they found that *polarization*, i.e., more extreme behavior with repeated intent measurement, also diminished with prior experience. These findings are consistent with more than one theoretical

explanation, including increased attitude accessibility, increased response fluency, and making inferences. It can be argued that experienced consumers are affected less because they possess: more accessible attitudes, greater cognitive fluency, and a large store of knowledge about the firm, respectively, prior to questioning, than those who are inexperienced. Although not specifically addressing this issue, Dholakia and Morwitz's (2002) study provides conflicting findings. Their study, conducted exclusively with experienced customers, still found a strong and persistent MME.

Fitzsimons and Morwitz (1996) found that at the brand level, the MME was manifested in different choices for current car owners and for those who did not own a car. Car owners gravitated toward their existing brand whereas non-owners were more likely to buy prominent high market share brands. These findings are more clearly consistent with an increased attitude accessibility explanation: current owners knew more about their current brand which became more accessible afterward; in contrast, non-owners knew more about leading car brands.

The practical implication of this moderator is that through judicious sample selection and post-survey interventions, researchers may be able to diminish the effects of questioning on their respondent base. For example, to minimize impact, the sample for a satisfaction survey could be over-weighted with experienced customers.

The customer's experience with the firm. Dholakia et al. (2004) found evidence of moderation by the customer's experience with the firm in their study. Note that this variable is conceptually different from the customer's experience with the product category because it measures the customer's knowledge of, and relationship with, the specific firm conducting the research. They found that the processes by which the MME occurred varied for novice and experienced customers. Novice customers were more susceptible to the *positivity effect*: they

made positive inferences regarding the firm based on the survey, which influenced their behavior. In contrast, experienced customers were influenced by increased accessibility of responses to survey questions which led them to behave in accordance with their expressed responses. Those who expressed satisfaction in the survey purchased more, whereas those who were dissatisfied purchased less than the control group. Thus, prior experience with the firm shifted the psychological process through which the MME occurred.

Although no studies have examined this issue, it is also likely that prior experiences with the firm could affect the types of inferences customers make regarding the survey. For example, customers who have had prior negative experiences may be much more skeptical and infer negative reasons for the survey than customers who have had positive experiences. More research is needed to understand the role played by firm experience on the MME.

***Respondent characteristics*.** The notion that individuals should be differentially susceptible to the MME based on their demographics and traits is intuitively appealing and of potential practical significance. Surprisingly, few studies thus far have sought to uncover individual differences in respondents' susceptibility to the MME. One exception is Borle et al. (2007), who studied this issue. Their empirically-oriented study was done in cooperation with a leading US-based automotive services store chain. In their paper, they developed a joint model of four customer behaviors during each service visit: (1) number of promotions redeemed; (2) number of services purchased; (3) time since the last visit in days; and (4) amount spent. They considered a number of customer characteristics as predictors: gender, age, tenure with the firm, the vehicle's manufacture year, median household income, and household size.

Borle et al. (2007) found a number of interesting moderating effects of customer characteristics on the MME. The effects of survey participation diminished with increasing age,

greater customer tenure, and with increasing age of the customer's vehicle. They argued that both age and tenure are indicative of customer experience and these customers are less likely to gain additional useful information from the survey, or to form measurement-induced judgments. In contrast, younger and newer customers are likely to have uncrystallized opinions regarding the firm, and the survey should impact them to a greater degree. Both household income and size also strengthened the MME for some of the behavioral variables. Interestingly, the customer's gender was the only characteristic studied that did not play a moderating role for any of the behaviors. The Borle et al (2007) study did not investigate the psychological reasons for the differences which remains an interesting and practically important issue to be studied.

***Firm characteristics*.** Borle et al. (2007) also examined the moderating role played by store-specific variables in influencing the MME's strength. The store-level variables studied were: (a) whether the store was company-owned or franchisee-owned; (b) whether it had a customer lounge; (c) its number of service bays; and (d) a measure of throughput times. The results revealed that survey participation had more beneficial effects on customers purchasing at company-owned stores than at franchisee-owned stores. In explaining this result, Borle et al. (2007) suggested that the difference could have arisen because company-owned stores offered a larger menu of services when compared to franchisee-stores. Consequently, after survey participation, customers visiting company-owned stores would have more opportunities to act in accordance with their positive evaluations than those visiting franchisee-owned stores. They also noted the possibility that employees at company stores might be more responsive, leading to more positive behaviors. This was the first, and to my knowledge, the only study to date, documenting the moderating role of firm characteristics on the MME.

The two moderators, customer characteristics and firm characteristics, support the

intuitive yet intriguing possibility that survey participants are differentially affected by the MME. There are a variety of trait variables that have the potential to play significant moderating roles. For example, the need for cognition (Cacioppo et al. 1996), the need to evaluate (Jarvis and Petty 1996), and the conscientiousness factor of the Big-Five traits (Conner and Abraham 2001) are good starting points to examine trait moderators for both lab- and field-based MME.

Behavior characteristics that increase ease of representation. Levav and Fitzsimons (2006) tested the moderating role of three aspects that increase the ease of representing the behavior. In one study, they manipulated *self-relevance of the intention question*, finding that when study participants were asked their own likelihood of flossing teeth (vs. the likelihood of one of their classmates), they were more susceptible to the MME. In another study, they manipulated *question frame*. Participants were either asked a straightforward, positively framed intent question about consuming fatty foods, the likelihood of not engaging in the behavior, or the likelihood of avoiding it completely. Compared to a control condition, all three groups consumed fewer chocolate-chip cookies (vs. rice cakes); additionally, those in the avoidance condition ate fewer cookies than either the intent or negation conditions. Levav and Fitzsimons (2006) argued this was because congruence between the negative attitude and avoidant behavior made the behavior easy to represent, increasing the MME's strength.

In their last study, the authors manipulated *congruence between regularity of target behavior and the frequency with which it was referenced* in the question. For regular behavior such as flossing teeth, the strength of the MME was greater when the question referenced a regular frequency (e.g., 7 times in the coming week) than an irregular frequency (e.g., 8 times in the coming week). No such moderating effect of behavioral frequency was found for reading for pleasure, a behavior that is usually performed irregularly. The authors argued that frequency

regularity positively affected the individual's ease of representation for regularly occurring target behaviors, but not for irregularly occurring behaviors.

Summary of mere measurement effect research

This review makes it clear that the mere measurement effect is a robust effect, replicated by a multitude of researchers in diverse settings, and influences consumer behavior significantly. However, unlike self-prophecy research, we don't yet know the magnitude of effect sizes across studies or the drivers of effect sizes. Additionally, two starkly different approaches to studying the effect can be discerned in the literature: studies conducted in the laboratory and field-based studies. There are significant differences between the two approaches in the study environment, types of participants, stimuli, and behaviors studied, all of which contribute to different theoretical explanations, boundary conditions, and practical implications. Researchers are advised to clearly define the approach chosen during study design, and test the current implicit but untested assumption that findings from one domain, say, the lab, apply to the other domain, the field.

THE SELF-PROPHECY EFFECT

Introduction to the self-prophecy effect research stream

Unlike mere measurement research which focuses on consumers' purchase and relational behaviors, research on the self-prophecy effect (SPE) exclusively examines effects of questioning on socially normative behaviors[2] and is defined as follows: "Asking people to make a self-prediction regarding a socially normative behavior influences the performance of that behavior in the future" (Spangenberg and Sprott 2006, p. 550). Tracing its origins more directly

[2] Of course, some purchase behaviors can be socially normative, for example, purchasing a gas-guzzler, sex toys, anti-depression medications, etc. In such cases, the self-prophecy effect applies.

to Sherman's (1980) study than the MME, the SPE has been documented in a wide range of socially normative contexts (see Sprott et al. 2006, for a recent review, and Spangenberg and Greenwald 1999, for an earlier review).

Studies have found self-prophecy to increase voter turnout in elections (Goldstein et al. 2008; Greenwald et al. 1987), increase attendance at health clubs (Spangenberg 1997), raise the commitment to a voluntary assessment of one's health and fitness (Sprott et al. 2004), increase recycling of aluminum cans (Spangenberg et al. 2003; Sprott, Spangenberg and Perkins 1999), reduce gender stereotyping (Spangenberg and Greenwald 1999), increase alumni donations to one's alma mater (Obermiller and Spangenberg 2000), and increase frequency of choosing low-fat snacks over less healthier options (Sprott, Spangenberg and Fisher 2003). These studies show that after self-prediction, behavior of respondents is biased in the socially normative direction. Questioning increases performance of socially desirable actions, and reduces performance of socially undesirable ones.

However, recent research studying the effects of questioning on risky behaviors specifically by children and adolescents has found normatively inconsistent behavioral effects (see Fitzsimons and Moore 2008 for a review). Risky behaviors are negatively-valenced from a normative standpoint, including such actions as drinking alcohol, having unsafe or unprotected sex, and using drugs. They are risky because their performance entails a threat to the mental and/or physical well-being of the respondent or others, immediately or in the future. For such behaviors, a number of recent studies reviewed by Fitzsimons and Moore (2008), have found that questioning adolescents increases behavior even when they report negative attitudes towards it.

Review of self-prophecy effect research studies

In one of the earliest demonstrations of the SPE which was positioned as a replication of

Sherman's (1980) study, Greenwald, Carnot, Beach and Young (1987) studied the effects of asking students by telephone to predict whether they would either register to vote, or actually vote. When compared to control groups who were not asked about performing these behaviors, there was an approximately 10% increase in the probability of voting registration and an approximately 25% increase in the likelihood of voting among those who made self-predictions.

In another early empirical demonstration in a consumer context, where the term "the self-prophecy effect" was first introduced, Spangenberg (1997) used a brief telephone survey and either asked members of an athletic health club who had not attended for at least one month whether or not they would visit the club in the future or did not ask this self-prediction question. During the subsequent six month period, the questioned group (who retained their membership for the duration) attended the club significantly more times than the control group.

Instead of individualized administration of prediction requests, Spangenberg and colleagues (2003) conducted an advertising campaign. They posted the question "Ask Yourself. Will You Recycle?" on a large (2 x 7 feet) electronic board, on actual-sized wooden stop signs at key entrances, and on flyers hung on bulletin boards in each classroom within a large building on a university campus. They measured recycling prior to this advertising campaign, as well as during and afterwards, by counting the proportion of aluminum cans purchased from vending machines in the building that were placed in recycling bins. They found that the campaign led to an increase in recycling behavior from 16% to 28%, that is, by 75%. This is one of the few QBE studies demonstrating the effect even when questions are neither posed nor answered individually by respondents. In this case, the question was asked through a mass-communication medium, and presumably answered by individuals internally when they encountered it. This finding significantly increases the scope of SPE effects, extending how they can be used by

practitioners interested in influencing socially normative behaviors of consumers.

A number of other studies have demonstrated occurrence of the self-prophecy effect in the lab and the field. Spangenberg and Greenwald (1999) examined effects of self-prediction on gender stereotype activation in a name-generation task (deciding whether famous individuals are male or female given their last names). Their results revealed that experimental group participants who were asked to predict whether they would be more likely to guess male names, female names, or both equally when they did not know the correct name, subsequently were more likely to guess female names erroneously than a control group. Self-prediction thus reduced the expression of gender stereotypes in this study.

Although positioned as a mere measurement effect study[3], Godin, Sheeran, Conner and Germain (2008) found that asking recent blood donors whether they would donate blood again in the next six months led to significantly greater registrations at blood drives as well as more successful donations six months as well as a year after questioning. The authors did not study underlying processes for these results. Sandberg and Conner (2009) extended these findings to the case of cervical screening among UK women, finding that asking about behavioral intentions plus anticipated regret from not performing the behavior increased attendance rates significantly (65%) when compared to asking the intentions question alone (44%). Both groups had higher attendance rates compared to a control group. The authors argued that higher levels of anticipated regret may bind people to their intentions and increase likelihood of behavior because failing to act would be associated with aversive affect.

By asking people whether or not they would donate money, Obermiller and Spangenberg (2000) were able to increase the rate of donation success from 30.4% to 49%. Instead of focusing

[3] This study is classified as a self-prophecy effect study here because of the socially normative nature of the examined behaviors.

on whether a question is asked, Liu and Aaker (2008) examined effects of the question's content. In studying consumers' willingness to give to charitable causes, they found evidence for a "time-ask effect" whereby asking consumers whether they would like to volunteer time to a charity versus asking whether they would like to donate money, or not asking any intent question at all, led to greater levels of monetary contributions. They explained the effect due to mindsets activated by the initial mention of time versus money. Answering a question about volunteering time increased salience of the action's (giving to charity) emotional significance for respondents, who viewed the charity as a means toward their happiness. This led to a more positive inclination toward giving to charity, and an increase in actual dollar contributions. This study provides a promising new direction to extend the scope of question-behavior effects, to understand how question content triggers processes influencing behaviors (see Bradburn et al. 1987, and Schwartz 1999, for detailed discussions regarding the effects of question content on responses given by survey participants). Relatedly, Gollwitzer and Oettingen (2008) have proposed that whether the question targets the critical behavior directly or indirectly will determine how the effect unfolds.

Similarly, Stutzer, Goette and Zehnder (2007) found that asking individuals to make a "strong active decision", i.e., articulate whether they would be willing to donate blood at one of a selection of specific dates and times, increased the probability of donating blood by 8.7% relative to a control group. Goldstein and colleagues (2008) compared the efficiency of measuring behavioral intentions vs. explicitly forming implementation intentions (Gollwitzer 1999) on enactment of both one-shot goals (voting on election day) and open-ended goals (voting early) either in the short-term (days) or long-term (months before). The authors found that intention measurement increased voter turnout for open-ended goals and for nearer one-shot goals, but not

for distant one-shot goals. Implementation intentions, on the other hand, were efficacious for both goal types over both lengths of time.

Spangenberg and Sprott (2006) found results consistent with the SPE, additionally demonstrating the moderating role of self-monitoring (discussed in detail in the next section). As this review makes clear, the self-prophecy effect finding behavioral changes in socially normative directions is robust and has received wide support. In a meta-analysis of published and unpublished studies (through 1999), Spangenberg and Greenwald (1999) found an average effect size of .19, with a range of .08 to .40. Including only SPE studies involving health-related behaviors, Sprott and Spangenberg (2006) reported an average effect size of .265.

SPE studies involving risky behaviors by children and adolescents. A number of recent studies have examined effects of questioning of adolescents on their risky behaviors. I include this line of research within the domain of self-prophecy effects because all the behaviors studied in this line, as well as the explanation advanced, have a significant socially normative component. Williams, Block, and Fitzsimons (2006) studied whether asking undergraduates questions about future drug use changed self-reports of actual drug use. Relative to a control group who was asked about exercising, the experimental group who was asked to make a prediction regarding how often they would use drugs reported more drug usage (2.8 times vs. 1 time) two months later.

Fitzsimons, Nunes and Williams (2007) examined the process by which these effects occur. Using a response latency task, they found that participants who had been asked how many times they were likely to skip class were much faster at categorizing “skip class” as positive in a response latency task when compared to a control group. However, their explicit attitude toward skipping class was negative. In a second experiment, they found that asking participants either

about drinking more than two alcoholic drinks at one time in the coming week or about watching television instead of studying resulted in increased self-reported drinking (3.2 times vs. 1.2 times) and television watching (3.9 times vs. 2.7 times).

Fitzsimons, Block and Williams (2007) replicated these results by measuring actual respondent behavior. In one study, participants were asked how many times they were likely to be distracted from studying in the coming week or an unrelated control question. After a delay, participants were provided with an opportunity to sign up for going to four movie screenings within a single week during the semester, a behavior that would result in a substantial distraction from studying. Those questioned about this "vice" behavior were significantly more likely to sign up for the movie screenings (76.6% versus 53.1% in the control group). In another study, students who were asked how many classes they would miss during the semester did miss more classes (3.78 class sessions vs. 2.95 class sessions) than the control group.

What is striking about studies examining risky behaviors in adolescents is that they appear to be similar in all important respects to the self-prophecy studies that find normatively consistent results. The types of behaviors examined are similar, study participants are comprised of young individuals (usually undergraduates), and one can make a persuasive case that these individuals are likely to hold conflicting implicit and explicit attitudes for activities such as exercising, choosing low fat snacks, or recycling, just as they do for skipping class or using drugs. So it is surprising that the pattern of results obtained in this set of studies is diametrically opposite to the traditional self-prophecy research (see Gollwitzer and Oettingen [2008]; Schneider, Tahk and Krosnick [2007]; Sherman [2008]; Spangenberg, Greenwald and Sprott [2008], for recent critiques of this line of research).

The controversy and the conflicting findings make this an exciting area with a number of

opportunities to resolve inconsistencies, and advance our understanding of the self-prophecy effect. For instance, one explanation could be the manner in which responses are elicited. Most (but not all) self-prophecy studies utilize binary (yes/no) responses whereas the risky behavior studies usually assess frequency of future behavior. It could be that these different response formats favor occurrence of one or the other process (e.g., cognitive dissonance vs. implicit attitude activation) leading to opposite effects on behavior.

Theoretical explanations for the self-prophecy effect

Unlike the MME, SPE studies appear to tap into the same phenomenon irrespective of the methodological approach used. Regardless of whether studies are done in the lab or the field, there is one leading explanation for the traditional (normatively consistent) self-prophecy effect's occurrence: *cognitive dissonance*. In this section, I discuss cognitive dissonance first, followed by a brief consideration of other candidate explanations, and finally the account advanced for normatively inconsistent effects for risky behaviors.

Questioning produces cognitive dissonance. The leading explanation for self-prophecy effects is that they are a manifestation of cognitive dissonance. The dissonance–based view of self-prophecy holds that providing a self-prediction about one's future behavior increases both salience of social norms associated with the behavior and one's prior failure to perform the behavior in a socially normative manner. Stated differently, when the prediction request is answered, respondents simultaneously become cognizant of what they should do as well as what they have done or not done in the past. Assuming that these cognitions are inconsistent (e.g., "I know what a good, moral and competent person should do, but I have failed to do so in the past or haven't done it as often as I should have. Now that I have an opportunity, I will do what I should have done all along"), the self-prediction task directly confronts the individual's self-

concept as a moral, competent, and good person. Given that such a person should behave in line with social norms but has not always done so in the past, cognitive dissonance is elicited through self-prediction (Aronson 1992). Once activated, dissonance serves as an aversive state, motivating behavior in the direction of social norms (Spangenberg et al. 2003; Spangenberg and Sprott 2006). Although yet untested, it could also be that self-prediction increases availability of prior failed behavioral attempts (Folkes 1988) contributing to self-concept confrontation.

The cognitive dissonance explanation is distinct from the explanations advanced for either the lab-based or field-based MME. The conjunction of *normative beliefs*, "I *should* behave in a certain way," *self-concept*, "… because I am a competent and morally good person", and a *behavioral discrepancy*, "I have not behaved this way in the past" is essential to instigating post-questioning action according to a cognitive dissonance explanation (Perkins et al. 2008). None of these elements play a role in mere measurement effects because there is no normative influence on the behaviors involved. However, at an abstract level, it can be argued that the raised awareness of these different cognitive structures constitutes a form of inference-making from responding to questions. A second point worth noting is that although most of the existing studies have focused on social norms (however see Chandon et al. 2007, for an exception), an identical process would unfold for a behavior that is of *personal normative significance*, for example, accomplishing a personal goal such as a resolution that one has set for oneself. In such cases, even if there is no socially normative significance, one would still expect questioning to result in cognitive dissonance, and influence behavior in the direction of the personal norm held by the respondent. This possibility needs to be explored in greater depth.

A number of studies support operation of the cognitive dissonance-based process. Spangenberg et al. (2003, Studies 3A and 3B) found that study participants viewing an

advertisement containing a prediction request reported significantly lower levels of psychological discomfort after making a prediction about other people's behavior, compared with a group not making this latter prediction. Such a finding is in line with cognitive dissonance theory which posits that when a person encounters a threat to one's self-esteem, he or she will bolster self-evaluations through downward comparisons with others (e.g., Wills 1981). In their fourth study, Spangenberg and colleagues (2003) found that giving participants the opportunity to affirm values central to their self-concept through selecting core values (from a supplied list) reduced levels of psychological discomfort relative to those who were not given the self-affirmation opportunity. Many of the studies examining moderators of the SPE (discussed in the next section) also provide evidence for cognitive dissonance.

The operation of cognitive dissonance in the SPE's occurrence raises interesting questions. For instance, Sprott et al. (2006) observe that it is not yet clear whether the effect would occur for behaviors that have a physiological component (e.g., nicotine dependence) which inhibits the effect of social norms (e.g., smoking, drug addiction, etc.). Under such circumstances, individuals may not feel personally responsible for the aversive consequences of those actions, attributing them to uncontrollable physiological forces. Similarly, if perceived obligations, permissions from respected individuals or groups, or peer pressure dampen the force of cognitive dissonance, the self-prophecy effect may diminish. On the flip side, factors that amplify dissonance such as intensifying social norms should accentuate the SPE.

Questioning activates a normative social identity. Perkins and colleagues (2008) recently argued that the SPE can be explained by the activation of a relevant normative social identity by questioning. Defining social identities as self-definitions that incorporate knowledge about a particular group that an individual belongs to or identifies with (Brewer 1991), the

authors proposed that answering a self-prediction question about a normative behavior, say recycling , should result in the activation of a "recycling" social identity which directs behavior. Positioning this explanation as an alternative to the cognitive dissonance account, the authors argued that unlike dissonance which lowers self-esteem, activating a positive social identity should increase self-esteem. In two experiments, one where implicit attitude for recycling was measured, and in the other where initial self-esteem was manipulated, the authors found that making a self-prediction resulted in greater self-esteem, which is inconsistent with a cognitive dissonance explanation. Interestingly, the authors did not find differences in explicit recycling attitudes between conditions, and concluded that "attitude accessibility is not a compelling explanation for self-prophecy" (p. 446). As of now, it is unclear when this process underlies the SPE's occurrence instead of cognitive dissonance. One possibility is that for certain types of normative behaviors that are central to one's social identity due to various reasons such as past experiences or salient beliefs, the identity activation process comes into play.

Questioning heightens self-awareness. Closely related to cognitive dissonance theory, Spangenberg and Greenwald (1999) invoked self-awareness theory (Duval and Wicklund 1972) to explain self-prophecy effects. According to this theory, the presence of self-focusing stimuli heightens the individual's self-focused attention, producing a state of objective self-awareness that involves attention to discrepancies between actual and ideal selves. The negative affect resulting from the perception of this discrepancy, in turn, leads the person to attempt to reduce the discrepancy. Greenwald and Spangenberg (1999) note the parallels between cognitive dissonance theory and self-awareness theory, observing that the two are "probably not distinguishable" (p. 84) in explaining self-prophecy effects.

Questioning evokes socially acceptable scripts of behavior. This account for the SPE is

consistent with the behavioral simulation explanation for the MME, and was advanced by Spangenberg and Greenwald (1999). It is one area where the MME and the SPE intersect. This explanation posits that for socially desirable behaviors, even though the individual may have preexisting scripts for a target situation, answering the question of what one *will do* is likely to produce a socially acceptable script, which will be mentally rehearsed by the individual. In cases where a preexisting script is lacking, social desirability of the target action will still lead participants to predict performance of the more socially desirable option. In either case, when the time to enact behavior arrives after being questioned, the socially acceptable script is likely to be enacted even overriding the usual script in the former case.

Questioning increases effect of implicit positive attitudes. Fitzsimons and Moore (2008) have persuasively argued that individuals, particularly adolescents and children, hold complex and ambivalent attitudes towards risky behaviors. As a result, the process underlying the effects of questioning on risky behaviors is distinct. At the conscious or explicit level, survey respondents are aware that the risky behavior could have negative consequences eventually, if not right away. Nevertheless, they are drawn to the behavior implicitly, at an unconscious level, often holding a positive implicit attitude toward its enactment. Emerging research (described earlier; e.g., Fitzsimons et al. 2007) indicates that in such cases, asking questions about risky behaviors toward which they hold both positive and negative attitudes often results in the positive, more implicit attitude guiding the actions of respondents and increasing enactment of risky behavior.

Despite its merits in accounting for the self-prophecy effect when the respondent's behavior after questioning conflicts with his or her explicit attitude regarding the behavior, it is worth noting this explanation is far from well-accepted (see Sherman 2008, for a detailed

critique). In addition to traditional self-prophecy effect studies, other research examining risky behaviors has found different results. For instance, in a study of playground behaviors among children aged 7-13 years, Morrongiello and Mark (2008) found that getting participants to advocate for safe-play behaviors while thinking about past failures to play safely on grounds resulted in reductions in self-reported risky play behaviors up to two months later. Thus, risky behavior change was in the socially normative direction. Other studies have found increased use of condoms among college students (Eitel and Friend 1999) and a greater propensity to obey speed limits (Fointiat 2004). Sherman (2008) describes other studies which show risky behaviors among adolescents after questioning.

These conflicting results suggest that our current state of knowledge regarding the conditions under which respondents behave in normatively consistent ways after answering questions, and when they behave in opposite ways is incomplete. Rather than categorizing an entire demographic group of individuals, i.e., adolescents, as having positive implicit attitudes toward risky behaviors, it might be productive in future research to identify the conditions that foster presence of conflicting implicit and explicit attitudes.

Moderators of the self-prophecy effect

Several studies have examined boundary conditions for the self-prophecy effect's occurrence. The roles played by normative beliefs of participants, their self-monitoring levels, the manner in which the self-prediction task is performed, respondent characteristics such as gender, and debiasing through providing advance warning have all been studied, and are discussed here.

***Normative beliefs of participants*.** Sprott, Spangenberg and Fisher (2003) argued that the individual's normative beliefs regarding the target behavior, defined as beliefs relating to what is

socially desirable or appropriate to do, are critical in the SPE's occurrence. They hypothesized that normative beliefs would act as a moderator of the SPE such that people with more strongly held normative beliefs would be more likely to exhibit a SPE when compared to those with a weaker normative stance on the issue. In two studies testing consumption of low-fat versus regular fat snacks, and participating in an assessment of one's health and fitness, they found support for their hypothesis. In summarizing their results, Sprott et al. (2003) noted:

> "The current research provides evidence that the self-prophecy effects appears to operate best when people possess strong beliefs about what is normatively right or wrong. Consequently, asking people to make predictions that are counter to these beliefs is unlikely to be effective. Indeed, because the self-prophecy effect appears to be driven by people's personal beliefs about what is appropriate, the most fundamental requirement for self-prophecy to manifest is a population (or subset thereof) that shares such beliefs. For example, a prediction request will not likely change the behaviors of heavy smokers, people who often litter, nonvoters, and those who do not engage regularly in exercise unless they become convinced that their current lifestyle with regard to these activities is inappropriate." (p. 429).

An interesting extension to this research would be to examine the interactions between personal beliefs and normative beliefs, and the conditions where they mutually enforce and contradict each other in contributing to the SPE.

Levels of self-monitoring. Spangenberg and Sprott (2006) studied the moderating role of an individual's level of self-monitoring on the SPE. Self-monitoring refers to the relative extent to which the individual's behavior is influenced by dispositional versus situational factors. For high self-monitors, behavior is influenced to a greater degree by situational factors, whereas the behavior of low self-monitors is influenced to a greater extent by dispositional factors. Low self-monitors are also influenced by messages appealing to their values (i.e., attitudes serving a value-expressive function) whereas high self-monitors are influenced via appeals to their status (i.e., attitudes serving a social-adjustive function).

The authors argued that the process of confronting the discrepancy between values and

prior action and reducing it through action should be more effective on low self-monitors because their attitudes are generally based on values, in contrast to high self-monitors, who base behaviors and attitudes mainly on situational factors. Consequently, they hypothesized that level of self- monitoring should act as a moderator of the SPE such that it is more likely to influence low versus high self-monitors. Two lab-based studies involving participation in a fifteen minute free health and fitness assessment, and donating a few hours of time to the American Cancer Society, provided support to this moderation hypothesis.

Characteristics of the self-prediction task. It appears that the manner in which the self-prediction task is performed also affects the SPE's strength. Both specificity and degree of cognitive elaboration have been shown to act as moderators. Sprott, Smith, Spangenberg, and Freson (2004) examined the role played by specificity of the prediction request. Their study involved participation in an assessment of one's overall physical fitness and health. In the specific prediction condition, participants were asked to predict whether they would participate in a voluntary health and fitness assessment, whereas the general prediction question asked about participation in one or more health and fitness activities. Their results revealed that whereas no one from the control or general prediction groups signed up for the assessment, 17.4% of those in the specific prediction condition signed up. The authors concluded that a specific prediction request is more likely to elicit a SPE than a general prediction request.

Van Kerckhove, Geuens, and Vermeir (2009) showed that participants with well-formed, i.e., cognitively elaborated self-predictions engaged in the focal behavior of choosing environmentally friendly products to a greater extent when compared to those who had ill-formed self-predictions, i.e., they were distracted during the self-prediction task.

Respondent characteristics. There is also some (although mixed) evidence from extant

studies that the respondent's gender can play a role in the SPE's occurrence. Spangenberg and Greenwald (1999, Experiment 1) found that for implicit gender stereotyping (a greater likelihood of judging famous individuals as male rather than female given their last names), males were less likely to stereotype when they were asked to predict what they would do relative to a control group. Females did not show this effect. However, they did not find this effect of gender in another study. More research is needed to better understand individual factors (e.g., including gender) that affect sensitivity to the SPE.

***Debiasing by providing advance warning*.** Williams, Fitzsimons, and Block (2004) showed that warning respondents in advance that responding to questions can change behavior dramatically reduced the question's impact on behavior. Respondents became more wary of the questioner, were more likely to perceive the question as an attempt to persuade them, and were able to consciously correct for the influence of the question. However, simply instructing respondents to think more deeply about questions, without an explicit warning that they may be attempts to influence, appears to have a counterproductive effect. In a study of voting behavior, Fitzsimons and Shiv (2001) instructed respondents to reflect about the question they were being asked. In this case, instead of diminishing impact, those asked to think more deeply actually showed a stronger effect of being asked the question versus a control group who was not given this instruction. These findings recommend that if the researcher's goal is to minimize the QBE's effect on respondents, warnings about the effects of questions should be given in advance, and should spell out their specific effects, rather than just providing broad instructions to think deeply.

Summary of self-prophecy effect research

Based on this review, a large number of studies have found evidence for the self-

prophecy effect in socially normative directions. Upon being asked to make a self-prediction, individuals realize they have not been acting as they should have, experience cognitive dissonance, and change their actions, behaving in line with social norms. The corpus of SPE findings make it clear that such a cognitive dissonance-based explanation applies to laboratory as well as field-based self-prophecy studies. Recent findings on normatively inconsistent behavior due to the clash of positive implicit and negative explicit attitudes regarding risky behaviors are intriguing, but they raise more questions than provide answers at this point.

PRACTICAL IMPLICATIONS OF QUESTION-BEHAVIOR EFFECT RESEARCH

Researchers of all social science disciplines rely on questioning to gather data, and are affected by the question-behavior effect. In this section, I discuss its significant practical implications including the assumption of measurement separability, assessing the net impact of survey research, using self-prophecy to encourage socially desirable behaviors, and understanding the implications of asking questions regarding risky behaviors to teenagers. The discussion is organized by implications of the MME, followed by implications of the SPE.

Implications of mere measurement effect research

Questioning the assumption of measurement separability. When conducting surveys, marketing researchers routinely make the assumption of *measurement separability* (i.e., the survey process only elicits existing opinions of consumers); it does not *form* or *influence* their opinions or behaviors in any way. In fact, segregating consumer research and political polling from sales or persuasion activities has always been one of the holiest grails of marketing research practice. Any attempt to combine these activities not only raises ethical concerns (for instance, by violating the code of ethics of the American Marketing Association), but attempts to do so

have been outlawed since the 1990s by legislators in the United States, through laws such as the Telemarketing Consumer Fraud and Abuse Prevention Act. The scope and range of findings regarding the MME and the SPE discussed in this chapter invalidate the assumption of measurement separability. Even when marketing researchers are ethical in their intent and execution, and have no interest in influencing respondents, they still end up having multi-faceted and long-term effects on consumers (see Machlin and Fitzsimons [2005] for a detailed discussion of this issue).

At first blush, the findings discussed here appear to bode well for certain types of consumer research activities such as customer satisfaction measurement. For instance, findings on the positivity effect imply that even customers expressing dissatisfaction will behave more relationally towards the firm subsequently (e.g., Dholakia et al. 2004). Likewise, long-term studies of survey effects usually find overall positive effects (Chandon et al. 2004; Dholakia and Morwitz 2002). However, these findings also raise thorny questions not only about the ethics of consumer research but also regarding the accuracy of forecasting models run using survey data (Chandon, Morwitz and Reinartz 2005; Heij and Franses 2006; Morwitz 2005). A number of unanswered questions emerge because of the QBE: Through what means is it possible to minimize or reverse the effects of questioning on consumers? In what way should anti-sugging laws be modified to reflect reality? How best should practices for conducting survey research promoted by the DMA, the CASRO, and other industry groups acknowledge occurrence of the QBE? By and large, academic researchers have bypassed these issues thus far. Practitioners, too, have not paid much attention to these questions.

However, recent studies have suggested ways to correct for mere measurement effects in forecasting models. Chandon, Morwitz and Reinartz (2005) proposed the following three-stage

procedure. In the first stage, available descriptive variables from surveyed customers, such as their demographic characteristics, are used to predict the pre-survey latent (unobserved) purchase intentions of both surveyed and non-surveyed consumers. In the second stage, the strength of the association between the pre-survey latent intentions and the post-survey behavior for the two groups is compared to assess the MME's strength. In the third stage, conversion schemes[4] (Jamieson and Bass 1989) used for forecasting are adjusted to account for the measurement effect depending on its cause and strength.

Heij and Franses (2006) extended this approach to directly predict purchase behaviors. Purchase was operationalized as a binary (yes/no) outcome. In their analysis of easy-to-prepare food products in two supermarkets, they found that forecasting models that neglected either the binary character of the data or the endogeneity of the measured intentions tended to under-estimate the MME. Despite these exceptional studies, designing a satisfactory strategy to reduce influences on behavior and errors in predictions, and to effectively communicate how these effects will influence different constituents such as consumers, managers, financial analysts, etc. remain open questions. In a broad sense, the implication of the QBE is that managers should be wary and cautious when conducting, interpreting, and using findings of consumer research studies. In a specific sense, they have to somehow adjust for the MME's occurrence when predicting what the broader population of (unsurveyed) consumers will do.

Assessing the net impact of conducting survey research for organizations. Firms usually treat the cost of conducting survey-based research as an expense. The conventional wisdom is that survey expenditures are recouped when useful information obtained from

[4] Conversion schemes are specific rules to convert purchase intentions into predictions of behavior (Jamieson and Bass 1989). For example, a five-response category purchase intention question may result in the use of a "75%-25%-10%-5%-2%" scheme, which specifies that 75% of consumers who stated that they would "definitely buy" (top box) will do so, 25% who stated they would "probably buy" will do so, and so on.

respondents enables more informed decision making. The QBE challenges this fundamental assumption by showing that survey-based research can produce direct and measurable benefits from changed respondent behavior. Specifically, because customers participating in surveys engage in greater purchase behavior in broad-based ways and over extended periods of time compared to non-participants, the firm generates an incremental revenue stream that can exceed the survey's cost.

The firm also accrues other substantial benefits. For example, a number of studies have shown that customers participating in surveys are more responsive to the firm's promotional efforts, increasing the efficacy of its other marketing programs (Borle et al. 2007; Dholakia et al. 2009). Positivity effect research has shown that survey participation generates long-lasting positive inferences regarding the firm (Dholakia et al. 2004). Past research has also shown that survey participants are less likely to defect to competitors (Dholakia and Morwitz 2002), resulting in yet a third incremental revenue stream to the firm sponsoring the survey. All these outcomes are quantifiable, and their combined economic effects may offset much if not all of the cost of the survey itself. In many cases, the net impact of conducting a survey will be positive. Alternatively, the revenue streams may be used to determine the return on investment of the research expenditure.

Take the case of one particular large firm analyzed by Dholakia et al. (2004). Although some numbers in this illustration were disguised for purposes of confidentiality, this analysis is instructive in how to assess the net impact of a survey after accounting for the MME. Their approach to approximating the value of the customer satisfaction measurement program to the firm had two elements on account of survey participation: (1) the reduction in customer defection, and (2) the increase in annualized customer spending.

The benefits of defection reduction were decomposed into: (a) more customers subsequently present within the firm's customer base, and (b) a lengthened average customer life span. Dholakia et al. (2004) estimated the increase in customers owing to the survey program by multiplying the number of customers participating by their higher retention rate. They estimated the lengthened average customer life span using the simple identity relating customer turnover and average life span, namely, that the inverse of the annual defection rate of a group of customers is equal to their average lifespan. Although this operation results in a substantial increase in years of additional customer life for survey participants, for purposes of the illustration they ignored the effect of increased customer tenure, focusing only upon the benefits realized in the first year following survey participation.

The final element of their calculation of economic impact was the increase in average spending by survey participants, which they obtained directly from the data in their study. In applying the calculations to the firm's data, they disguised the survey data for purposes of confidentiality by multiplying observed values by arbitrary constants. Also for reasons of confidentiality, the economic impact estimation results are illustrated for three different assumed annual rates of company-wide customer defection (10%, 15%, and 20%), rather than reveal the firm's actual customer defection rate. Note that such an analysis is conservative because it did not include benefits such as increased responsiveness to the firm's promotions, and greater affinity to the brand due to positive inferences. As can be seen from these analyses which are summarized in Table 4, the economic impact of the survey program was significant and positive for all three levels of customer defection rates considered.

[Insert Table 4 about here]

As this illustration indicates, conducting a satisfaction survey could have net positive

impact for many firms. In such cases, the more accurate way of thinking about survey-based marketing research is that it is an investment in strengthening the firm's relationship with customers rather than a cost to gather customer opinions. Nevertheless, the limits of these positive effects remain presently unknown. It is still not clear how much research is too much, or whether there are certain conditions under which the effects from survey participation can turn negative. Research examining effects of repeated satisfaction survey participation on customer behavior as well as surveying different proportions of the customer population is sorely needed to help answer these questions.

Implications of self-prophecy effect research

Using the self-prophecy effect to encourage socially desirable behavior. Getting an individual to provide a self-prediction regarding his or her future behavior is a powerful social influence technique (Spangenberg and Greenwald 2001). As studies have shown, this approach is effective for a variety of normative behaviors and may equally apply to other behaviors not yet examined such as lowering personal debt and increasing saving, losing excess body weight, and practicing environmentally conservative behaviors. The finding that mass-communicated self-prediction requests are effective extends how the SPE can be utilized by practitioners. As one example, Spangenberg and colleagues (2003) suggested that with the help of advertisements, the sportswear retailer Patagonia might be able to increase sales of its eco-friendly products by asking its target consumers to make self-predictions regarding their support of environmentally friendly firms.

It is noteworthy that despite its practical significance, the effects of mass-communicated questioning have not yet been studied by MME researchers. For instance, instead of conducting costly telephone surveys, one can easily envision placing colorful post-card mailers or banners in

retail stores, boldly lettered with the question "Ask Yourself: Are You Satisfied With Our Service Today?" obtaining the MME via a more cost-effective approach. Research studying whether mass-communicated questioning generalizes to field-based MME, and if so, under what circumstances, deserves further attention from QBE scholars.

Implications of asking questions regarding risky behaviors to teenagers. Fitzsimons and Moore (2008) argued that in a public health context, screening adolescents for risky behaviors such as drug use, alcohol drinking, unsafe sex, etc. could potentially increase their subsequent levels of such behaviors. They noted that screening and surveillance of teenagers is widely practiced with tens of thousands of teenagers being questioned by various governmental and private organizations on an annual basis. Whereas individual screening assesses adolescents on a one-on-one basis for risky behaviors when they come in contact with the medical system (e.g., when they visit the doctor), population-based surveillance refers to large scale survey-based studies conducted to determine the prevalence of health risk behaviors in the population at large. Fitzsimons and Moore (2008) note that these procedures largely ignore the possibility of occurrence of the QBE or its potential consequences. Consequently, they observe that "Instead of being purely innocuous information gathering tools, surveillance surveys have a potential downside risk which is quite dangerous if not followed up by more thorough interventions. These questions function as form of influence that operates largely outside the respondent's ability to detect or correct" (p. 90).

To counteract the potentially serious effects of questioning on children and adolescents, public health and medical officials may use one of two approaches when asking these consumer groups about their risky behaviors: (1) follow up the questioning with an appropriate intervention, or (2) exercise care, and utilize known findings to ask minimally influential

questions. Based on the fact that the goal of individual screening is to identify risky behaviors and treat them, Fitzsimons and Moore (2008) observe that the earlier the medical professionals intervene, the better and cheaper will be the outcome. They offer several suggestions for effective interventions, including an emphasis on preserving and respecting adolescent autonomy, avoidance of categorical rules or statements about not engaging in risky behaviors, and their implementation in schools or communities after they have been tested.

In cases where providing interventions after screening is not possible, such as for large-scale surveillance studies, Fitzsimons and Moore (2008) suggest altering the survey itself to provide advance warnings, having the respondent commit to not engaging in the behavior, and changing the target and framing of the question, all of which have been shown by prior QBE research to minimize the effect of questioning on subsequent behaviors (e.g., Fitzsimons et al. 2007, Levav and Fitzsimons 2006). In summary, much research is needed before we can fully understand the scope and direction of the effects of questioning on risky behaviors of adolescents, and specific ways to minimize any adverse effects that occur.

CONCLUSION

The question-behavior effect challenges basic assumptions of researchers that use a questioning methodology to assess internal states of individuals and predict their future behavior. Over the last fifteen years or so, it has emerged to become a vibrant and impactful research area within marketing that is amenable to multiple perspectives and research emphases. During this time, research on the QBE has coalesced into two distinct streams, the mere measurement effect dealing with purchase behaviors without normative significance, and the self-prophecy effect concerned with socially normative behaviors.

In spite of the recent attempt at integration, throughout this critical review, I argued that there are fundamental differences between the two effects that preclude the complete merging of the areas. Not only are the types of behaviors studied different in the two cases, the underlying underlying processes, many of the boundary conditions, and the practical implications are unique. Whereas the primary process supporting occurrence of the SPE is cognitive dissonance, increased attitude accessibility is the most favored explanation for the lab-based MME, and the generation of inferences from survey participation supports the field-based MME's occurrence.

The moderators of the effects are also different. The consumer's experience level (in the product category and with the firm), firm characteristics such as whether the store is company- or franchisee-owned, and behavior characteristics facilitating representation of the behavior have been shown to moderate the MME, and participants' normative beliefs, levels of self-monitoring, and specificity of the prediction request are the SPE's known moderators. Respondent characteristics have been shown to moderate both effects, but even in this case, researchers have found some gender differences in occurrence of the SPE, but none for the MME. These differences are summarized in Table 2.

For the sake of conceptual clarity and to advance knowledge regarding the QBE most efficiently, it seems prudent to retain the distinct labels of the two effects, rather than abandoning them in favor of the common "Question-Behavior Effect" label. In fact, because of the distinctions, I suggested that researchers should clearly specify which area they are building on and contributing to when developing hypotheses, designing studies, and interpreting their results. Furthermore, even within each effect there are distinct streams (see Table 3) that are important to distinguish between in future research.

Although the two effects have largely distinct domains, there are some explanations for

mere measurement and self-prophecy which do overlap, the most notable of which is behavioral simulation. Any behavior, whether of socially normative significance or not may be simulated or imagined when the respondent has to answer questions regarding its future enactment, increasing its processing fluency, and enhancing likelihood of its subsequent enactment. Furthermore, although yet untested, it seems reasonable to expect that such a process should be insensitive to how questions are phrased, and by what means they are asked. Furthermore, there is recent evidence that repeated self-prediction increases strength of the self-prophecy effect (Van Kerckhove et al. 2009), analogous to the polarization effects sometimes found for the MME.

It is only when we consider what happens *in toto* when a question is asked that differences between the two effects begin to emerge. For instance, in addition to behavioral simulation and implementation intention formation, we may expect a motivational boost from cognitive dissonance to the extent that the behavior entailed is socially normative and the respondent perceives a discrepancy between his or her past behavior and normative beliefs. In contrast when personal interest is high or a long-standing relationship with the firm exists but the behavior has no normative significance, inference-making will likely contribute to the effect of questioning. As this discussion suggests, rather than the outcome of a single process, mere measurement and self-prophecy are each likely to be driven by a combination of psychological processes, which, depending on the individual and situation, interact in synergy or opposition, to influence behavior.

A common concern, applicable to both field-based mere measurement and self-prophecy research is the problem of *respondent self-selection*. According to this criticism, only certain types of respondents such as those with an avid interest in, or an extreme opinion regarding the focal issue, or those who generally like to answer surveys, agree to respond to the researcher's

questions. Consequently, the critics argue that the question-behavior effect is a manifestation of the atypical cognitive processing and behavior of these respondents rather than that of the entire relevant population. For instance, some critics suggest that because of the inherent differences, satisfaction survey respondents would have purchased more than the control group even in the absence of the question manipulation. This is a serious concern regarding much of QBE research and is applicable to all field-based studies[5]. One way that QBE researchers respond to the self-selection criticism is by trying to increase the survey's response rates. When higher response rates are achieved, they are used to argue that the self-selection bias is minimized or is less of a concern in their particular study. In many cases, QBE researchers do not consider the issue of self-selection bias at all.

Recent research by Anderson, Hansen and Tripathi (2009) specifically studied the impact of self-selection on the field-based MME. The authors provided a methodology that not only controls for *self-selection bias*, but also for *targeting bias*, which they defined as the possibility that the researcher targets surveys to particular non-representative sub-segments of individuals within the population. Intriguingly, their study conducted with customers of a direct marketing firm found that after controlling for the targeting and self-selection biases, the MME diminished to insignificance for both purchase frequency and spending behaviors of customers. Similarly, another recent study conducted by Algesheimer and colleagues (2009) found that when self-selection is taken into account and controlled for, previously strong positive behavioral effects of joining a customer community became insignificant for a majority of behaviors examined. Although prior QBE research has ignored or downplayed the seriousness of the self-selection

[5] Note that this criticism is not applicable to lab-based studies of either mere measurement or self-prophecy, because in these cases, all participants complete the study, and are randomly assigned to either the experimental or the control groups.

bias, Anderson, Hansen, and Tripathi's (2009) study forcefully underscores the need to better understand the implications of this bias for occurrence of the QBE. Methodologically, these findings point to the importance of minimizing such biases via appropriate sampling and survey administration procedures, and for controlling for them during analysis.

Finally, despite the fact that the first QBE study appeared nearly thirty years ago, and its benefits to firms and other organizations have been confirmed time and again across a body of studies, it is surprising how many practitioners such as marketing researchers, opinion pollsters, and public health officials still remain unaware of this effect. As researchers with a vested interest in seeing our work have tangible impact on the profession we serve, it is incumbent upon us to take the time and effort to publicize our findings and explain their value to practitioners at every available opportunity. In this quest for creating awareness and publicizing our research, mere measurement and self-prophecy researchers can stand shoulder-to-shoulder and speak with one voice.

TABLE 1
Summary of Distinctions Found Across Question-Behavior Effect Studies

Q-B research focus	Effects of: measuring purchase intentions (in lab and field), measuring satisfaction with firm-sponsored surveys, making self-predictions regarding socially desirable and undesirable behaviors, forewarning customers they will evaluate surveys through individualized questioning, screening and monitoring surveys on risky behaviors of children and adolescents, mass-communicated self-prediction requests through advertisements, asking hypothetical questions.
Type of question (independent measure) assessed	Behavioral intention (e.g., intention to purchase or intention to recycle), behavioral expectation, self-prediction using a dichotomous (yes/no) question, attitude (e.g., overall or transaction-specific customer satisfaction)
Type of behavior (dependent measure) assessed	One-time behavior (e.g., purchase of an item), behavioral pattern (e.g., frequency of unprotected sex or drug use), behavior repetition (Chandon et al. 2007), choice (e.g., purchase of one brand vs. another brand), self-reported behavior (Williams, Fitzsimons and Block 2004), written commitment to perform behavior in the future (Sprott, Spangenberg and Fisher 2003), behavioral expectation (Janiszewski and Chandon 2007)
Response Modality	Paper and pencil, telephone, face-to-face interview, mass communicated "ask yourself" advertisement, individual mailers
Methodology	Laboratory experiment, Controlled field experiment, Panel-based field data
Theoretical Explanations	Questioning: (1) increases accessibility of attitudes, (2) results in behavioral simulation and increases response fluency, (3) creates a perceptual image of action (ideomotor theory), (4) polarizes attitudes, (5) generates positive inferences, (6) generates a broad range of inferences, (7) produces cognitive dissonance, (8) heightens self-awareness, (9) evokes socially acceptable scripts of behavior, (10) increases the effect of implicit positive attitudes (even when explicit attitudes regarding behaviors are negative)
Study distinctions	(1) Whether the behavior studied is socially desirable or not, (2) How soon the behavior is assessed after questioning, (3) Whether the behavior is assessed in the laboratory in a controlled setting or whether it is assessed in the field, (4) How effortful, i.e., easy or difficult to implement, the behavior is.

TABLE 2
Comparative Summary of Question-Behavior Effect Research Streams

	Lab-Based Mere Measurement Effect Research	Field-Based Mere Measurement Effect Research	Normatively Consistent Self-Prophecy Effect Research	Normatively Inconsistent, Implicit Attitude-Driven Self-Prophecy Effect
Time-frame of Effect	Generally short; behavior is measured within minutes or hours of questioning	Generally long; behavior is measured weeks, months, even years after questioning	Studies have examined both short and long-term effects	Empirical findings are based mostly on short-term effects; conclusions have been applied to longer term
Methodology	Only lab studies	Only field experiments	Lab studies and field experiments	Lab studies and field experiments
Theoretical Explanation	Increased attitude accessibility, increased response fluency behavior simulation, ideomotor process, attitude polarization. Automatic processing plays a larger role.	Generation of positive inferences, generation of a broad range of inferences. Largely driven by deliberate information processing.	Cognitive dissonance, activation of normative social identity, increased self-awareness, script evocation. Driven mostly by deliberate processing.	Behavior influenced by positive implicit attitudes instead of negative explicit attitudes. Automatic processing counteracts controlled processing.
Questions Asked (Independent Measures)	Purchase intentions	Purchase intentions, customer satisfaction	Self-prediction by responding to "yes/no" question; likelihood of performing a behavior	Behavioral intentions regarding risky behavior
Effects Studied (Dependent Measures)	Self-reported behavior, actual behavior, choice	Purchase of new product, customer defection, customer profitability, number of services purchased, inter-purchase time	Actual behavior	Self-reported behavior, actual behavior
Behaviors/Products/ Industries Studied in Research	Candy bars, ice cream treats, usually cheap, frequently purchased food items	Automobiles, PCs, financial services, online grocery, automotive maintenance services (quick-lube oil change)	Recycling, voting, donating to charity, attending a health club, health and fitness screening, alumni donations, gender stereotyping, choosing a	Drug use, skipping classes, drinking alcohol, watching television instead of studying, class attendance

			low fat snack, attending cervical screening, donating blood	
Moderators of the effect	Experience with product, behavior characteristics that increase ease of representation	Experience with product, experience with firm, customer characteristics (demographics), firm characteristics	Normative beliefs of participants, self-monitoring levels, specificity of self-prediction request, respondent characteristics	Debiasing by providing advance warning
Notable Studies	Fitzsimons and Morwitz (1996), Fitzsimons and Williams (2000), Chapman (2001), Fitzsimons and Shiv (2001), Morwitz and Fitzsimons (2004), Levav and Fitzsimons (2006), Janiszewski and Chandon (2007)	Morwitz, Johnson and Schmittlein (1993), Ofir and Simonson (2001), Dholakia and Morwitz (2002), Chandon, Morwitz and Reinartz (2004, 2005), Dholakia, Morwitz and Westbrook (2004), Borle, et al (2007), Ofir, Simonson and Yoon (2009), Dholakia, Singh and Westbrook (2009)	Sherman (1980), Greenwald et al. (1987), Spangenberg (1997), Obermiller and Spangenberg (2000), Spangenberg and Greenwald (1999), Spangenberg et al. (2003), Sprott, Spangenberg and Fisher (2003), Sprott et al. (2004), Spangenberg and Sprott (2006), Stutzer, Goette and Zehnder (2007), Perkins et al. (2008), Godin et al. (2008), Goldstein et al (2008), Liu and Aaker (2008), Sandberg and Conner (2009)	Williams, Block and Fitzsimons (2006), Fitzsimons, Nunes and Williams (2007), Fitzsimons, Block and Williams (2007)

Table 3
Unanswered QBE Research Questions

Research opportunities regarding mere measurement effect

- In a meta-analysis of MME studies, what is the average effect size and what is the range of effect sizes across studies? Are there differences in effect sizes between lab-based and field-based MME studies? What are the other drivers of differences in effect sizes?
- Do mass-communicated requests produce the MME? When and to what degree?
- What aspects of lab-based MME findings can be extrapolated to field-based MME studies? What aspects are unique and not transferable between studies?
- How does behavior of customers who anticipate and later complete satisfaction surveys, customers who only complete satisfaction surveys, and a control group, differ from one another? Comparing these behaviors within a single study will help determine relative magnitudes of different effects on behavior.
- In studying how increased attitude accessibility contributes to the MME, does product category experience increase the relative role of automatic processing vs. effortful processing? If yes, to what degree?
- What role does increased attitude accessibility play in longer-term occurrence of the field-based MME? Morwitz and Fitzsimon's (2004) lab-based experimental paradigm could be extended to execute in the field over a longer time period.
- Does questioning facilitate formation of implementation intentions? Are certain types of questions (e.g., purchase intentions) more conducive to forming implementation intentions?
- Can processing fluency produce a MME for complex behavior involving self-regulation or for purchases of expensive products?
- Can processing fluency effects persist for weeks or months after initial exposure?
- What role does the ideomotor process play in the MME's occurrence?
- What role does attitude polarization play in the MME's occurrence? Are there certain circumstances or conditions under which it plays a stronger role?
- Under what circumstances (e.g., telephone-based vs. online, use of a professional interviewer, opportunity to provide open-ended feedback, follow up by firm to explain how customer feedback was used) do customers produce positive inferences from survey participation?
- What is the specific cognitive process through which the positivity inference process occurs?
- What other types of inferences can customers form because of survey participation? Under what circumstances do these inferences work in tandem, counteract, or are unrelated to one another?
- How do prior experiences with the firm affect the types of inferences made by customers regarding the survey?
- Which individual differences between respondents (e.g., demographics, traits) moderate the

MME's occurrence? What are the psychological reasons for these differences?

Research opportunities regarding self-prophecy effect

- How does question content influence occurrence of the SPE? Does questioning that targets the critical behavior directly or indirectly change the underlying process?
- Does self-prediction increase availability of prior failed attempts contributing to self-concept confrontation?
- Can personal norms (instead of social norms) contribute to cognitive dissonance for personally normative behaviors such as resolutions and generate the SPE?
- How do personal beliefs and normative beliefs interact in occurrence of the SPE? Under what conditions do they mutually reinforce and contradict each other?
- Does the SPE occur for behaviors that have a physiological component (e.g., nicotine dependence) which inhibits the effect of social norms?
- Do perceived obligations, permissions from respected individuals or groups, or peer pressure dampen the force of cognitive dissonance?
- Under what conditions does self-prediction activate a normative social identity instead of cognitive dissonance?
- Which individual factors affect sensitivity to the SPE, and through what means?
- Under what circumstances do explicit and implicit attitudes for socially normative behavior contradict each other? What moderating variables determine relative strength of the two attitudes?
- What can we learn from the self-regulation literature on how to enhance positive impact of the SPE and mitigate potentially negative impacts of positive implicit attitudes for risky behaviors?

Common research opportunities across mere measurement and self-prophecy research

- What proportion of the QBE (across studies) is due to respondent self-selection? What aspects of respondent self-selection are particularly contributory to the effect's occurrence? How can the researcher control for the selection bias?
- In what ways can the QBE be minimized or reversed? Does forewarning help minimize the effect, and under what circumstances?
- What are the effects of repeated questioning on different customer groups?
- In what ways should anti-sugging laws be modified to reflect reality? How best should practices for conducting survey research promoted by the DMA, the CASRO, and other industry groups acknowledge occurrence of the QBE?
- How can managers correct for errors in prediction arising from the QBE in their forecasting models?

Table 4
Illustrative Analysis to Demonstrate Economic Impact of Survey Participation

	Assumed Customer Defection Rate		
	10%	15%	20%
a. Number of satisfaction survey participants	374,920	374,920	374,920
b. Reduction in defection rate	62%	62%	62%
c. Saved customers (a * b)	23,160	34,740	46,320
d. Increase in customer lifespan (yrs.)	16.16	10.77	8.08
e. Increase in annual spending ($)	$12.23	$12.23	$12.23
f. Total gain in revenue, first year	$2,255,283	$3,382,925	$4,510,566
g. Annual survey expense	$2,049,675	$2,049,675	$2,049,675
h. Net gain from survey participation (f – g)	$205,609	$1,333,250	$2,460,892
i. "Return" on satisfaction survey program	10.0%	65.0%	120.1%

References

Aarts, Henk, Peter M. Gollwitzer and Ran R. Hassin (2004), "Goal Contagion: Perceiving is For Pursuing," *Journal of Personality and Social Psychology*, 87(1), 23-37.

Abelson, Robert P. (1981), "Psychological Status of the Script Concept," *American Psychologist*, 36(7), 715-729.

Alba, Joseph W., J. Wesley Hutchinson, and John G. Lynch, Jr. (1991), "Memory and Decision Making," In Thomas S. Robertson and Harold H. Kassarjian (Eds.), *Handbook of Consumer Behavior* (pp. 1-49), Englewood Cliffs, NJ: Prentice-Hall.

Anderson, Eric T., Karsten Hansen and Manish Tripathi (2009), "Measuring the Mere Measurement Effect in Non-Experimental Field Settings," *Working paper*, Northwestern University.

Aronson, Elliot (1992), "The Return of the Repressed: Dissonance Theory Makes a Comeback," *Psychological Inquiry*, 3(4), 303-311.

Bargh, John A., Peter M. Gollwitzer, Annette Lee-Chai, Kimberly Barndollar, and Roman Trötschel (2001), "The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals", *Journal of Personality and Social Psychology*, 81(6), 1014-1027.

Borle, Sharad, Utpal M. Dholakia, Siddharth S. Singh, and Robert A. Westbrook (2007), "The Impact of Survey Participation on Subsequent Customer Behavior: An Empirical Investigation," *Marketing Science*, 26(5), 711-726.

Bradburn, Norman M., Lance J. Rips and Steven K. Shevell (1987), "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys," *Science*, 236, 157-161.

Brewer, Marilynn B. (1991), "The Social Self: On Being the Same and Different at the Same Time," *Personality and Social Psychology Bulletin*, 17(5), 475-482.

Cacioppo, John T., Richard E. Petty, Jeffrey A. Feinstein, and W. Blair G. Jarvis (1996), "Dispositional Differences in Cognitive Motivation: The Life and Times of Individuals Varying in Need for Cognition," *Psychological Bulletin*, 119(2), 197-253.

Chandon, Pierre, Vicki G. Morwitz, Ronn J. Smith, Eric R. Spangenberg, and David E. Sprott (2007), "When Does the Past Repeat Itself? The Role of Behavior Prediction and Personal Norms," Working Paper, INSEAD.

Chandon, Pierre, Vicki G. Morwitz and Werner J. Reinartz (2004), "The Short- and Long-Term Effects of Measuring Intent to Repurchase," *Journal of Consumer Research*, 31 (December), 566-572.

Chandon, Pierre, Vicki G. Morwitz and Werner J. Reinartz (2005), "Do Intentions Really Predict Behavior? Self-Generated Validity Effects in Survey Research," *Journal of Marketing*, 69 (2), 1-14.

Chapman, Kenneth J. (2001), "Measuring Intent: There's Nothing 'Mere' about Mere Measurement Effects," *Psychology & Marketing*, 18 (8), 811-841.

Conner, Mark and Charles Abraham (2001), "Conscientiousness and the Theory of Planned Behavior: Toward a More Complete Model of the Antecedents of Intentions and Behavior," *Personality and Social Psychology Bulletin*, 27(11), 1547-1561.

Converse, Philip E. (1970), "Attitudes and Non-Attitudes: Continuation of a Dialogue," In Edward R. Tufte (Ed.), *The Quantitative Analysis of Social Problems* (pp. 168-189). Reading, MA: Addison-Wesley.

Dholakia, Utpal M. and Richard P. Bagozzi. (2003). "As time goes by: How goal and implementation intentions influence enactment of short-fuse behaviors," *Journal of Applied Social Psychology*, 33 (5), 889-922.

Dholakia, Utpal M. and Vicki G. Morwitz (2002), "The Scope and Persistence of Mere Measurement Effects: Evidence from a Field Study of Customer Satisfaction Measurement," *Journal of Consumer Research*, 29 (2), 159-167.

Dholakia, Utpal M., Vicki G. Morwitz and Robert A. Westbrook (2004), "Firm-Sponsored Satisfaction Surveys: Positivity Effects of Customer Purchase Behavior?" *MSI Reports*, 04-121.

Dholakia, Utpal M., Siddharth S. Singh, and Robert Westbrook (2009), "Understanding the Effects of Satisfaction Surveys on Customer Purchasing Behavior: Evidence From the Automotive Services Industry," *Working paper*, Rice University.

Duval, Shelley and Robert A. Wicklund (1972), *A Theory of Objective Self-Awareness*. New York: Academic Press.

Eitel Patricia and Ronald Friend (1999), "Reducing Denial and Sexual Risk Behaviors in College Students: A Comparison of a Cogitive and a Motivational Approach," *Annals of Behavioral Medicine*, 21(1), 12-19.

Fazio, Russell H., David M. Sanbonmatsu, Martha C. Powell, and Frank R. Kardes (1986), "On the Automatic Activation of Attitudes," *Journal of Personality and Social Psychology*, 50(2), 229-238.

Feldman, Jack M. and John G. Lynch, Jr. (1988), "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (3), 421-435.

Fitzsimons, Gavan J. and Sarah G. Moore (2008), "Should I Ask Our Children About Sex, Drugs, and Rock & Roll? Potentially Harmful Effects of Asking Questions About Risky Behaviors," *Journal of Consumer Psychology*, 18(2), 82-95.

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.

Fitzsimons, Gavan J., Joseph C. Nunes, and Patti Williams (2007), "License to Sin: The Liberating Role of Reporting Expectations," *Journal of Consumer Research*, 34 (1), 22-31.

Fitzsimons, Gavan J. and Baba Shiv (2001), "Nonconscious and Contaminative Effects of Hypothetical Questions on Subsequent Decision Making," *Journal of Consumer*

Research, 28(September), 224-238.

Fitzsimons, Gavan J., and Patti Williams (2000), "Asking Questions can Change Choice Behavior: Does it do so Automatically or Effortfully?" *Journal of Experimental Psychology: Applied*, 6 (3), 195-206.

Fointiat, Valerie (2004), "'I Know What I Have to do, But…' When Hypocrisy Leads to Behavioral Change," *Social Behavior and Personality*, 32(8), 741-746.

Folkes, Valerie S. (1988), "The Availability Heuristic and Perceived Risk," *Journal of Consumer Research*, 15(1), 13-23.

Folkes, Valerie S., and Vanessa M. Patrick (2003), "The Positivity Effect in Perceptions of Services: Seen One, Seen Them All?" *Journal of Consumer Research*, 30 (1), 125-137.

Gilbert, Daniel T., and Gregory J. Hixon (1991), "The Trouble of Thinking: Activation and Application of Stereotypic Beliefs," *Journal of Personality and Social Psychology*, 60(4), 509-517.

Godin, Gaston, Mark Conner, Paschal Sheeran, and Marc Germain (2008), "Asking Questions Changes Behavior: Mere Measurement Effects on Frequency of Blood Donation," *Health Psychology*, 27(2), 179-184.

Goldstein, Daniel G., Kosuke Imai, Anja S. Göritz and Peter M. Gollwitzer (2008), "Nudging Turnout: Mere Measurement and Implementation Planning of Intentions to Vote," *Working paper*, London Business School. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=977000

Gollwitzer, Peter M. (1999), "Implementation Intentions: Strong Effects of Simple Plans," *American Psychologist*, 54(7), 493-503.

Gollwitzer, Peter M., and Gabriele Oettingen (2008), "The Question-Behavior Effect From An Action Control Perspective," *Journal of Consumer Psychology*, 18(2), 107-110.

Greenwald, Anthony G., Catherine G. Carnot, Rebecca Beach, and Barbara Young (1987), "Increasing Voting Behavior by Asking People if They Expect to Vote," *Journal of Applied Psychology*, 72 (2), 315-318.

Hassin, Ran R., Henk Aarts, and Melissa J. Ferguson (2005), "Automatic Goal Inferences," *Journal of Experimental Social Psychology*, 41 (2), 129-140.

Heij, Christiaan and Philip Hans Franses (2006), "Prediction Beyond the Survey Sample: Correcting for Survey Effects on Consumer Decisions," *Econometric Institute Report EI 2006-48*, Erasmus University Rotterdam.

Jacoby, Larry L. (1991), "A Process Dissociation Framework: Separating Automatic From Intentional Uses of Memory," *Journal of Memory and Language*, 30 (5), 513-541.

James, William (1950), *The Principles of Psychology*, New York: Dover (Originally Published, 1890).

Jamieson, Linda F. and Frank M. Bass (1989), "Adjusting Stated Intention Measures to

Predict Trial Purchase of New Products: A Comparison of Models and Methods," *Journal of Marketing Research*, 26(August), 336-345.

Janiszewski, Chris and Elise Chandon (2007), "Transfer-Appropriate Processing, Response Fluency, and the Mere Measurement Effect," *Journal of Marketing Research*, 44 (2), 309-323.

Jarvis, W. Blair G., and Richard E. Petty (1996), "The Need to Evaluate," *Journal of Personality and Social Psychology*, 70(1), 172-194.

Johnson, Allison R., and Valerie S. Folkes (2007), "How Consumers' Assessments of the Difficulty of Manufacturing a Product Influence Quality Perceptions," *Journal of the Academy of Marketing Science*, 35(3), 317-328.

Kardes, Frank R., Chris T. Allen and Manuel J. Pontes (1993), "Effects of Multiple Measurement Operations on Consumer Judgment: Measurement Reliability or Reactivity?" In Leigh McAlister and Michael Rothschild (Eds.), *Advances in Consumer Research, Vol. 20* (pp. 280-283), Provo, UT: Association for Consumer Research.

Levav, Jonathan and Gavan J. Fitzsimmons (2006), "When Questions Change Behavior: The Role of Ease of Representation," *Psychological Science*, 17 (3), 207-213.

Liu, Wendy and Jennifer Aaker (2008), "The Happiness of Giving: The Time-Ask Effect," *Journal of Consumer Research*, 35 (October), 543-557.

Lusk, Jayson, Leatta McLaughlin, and Sara Jaeger (2007), "Strategy and response to purchase intention questions," *Marketing Letters*, 18(1), 31-44.

Machin, Jane E. and Gavan J. Fitzsimons (2005), "Marketing by Mistake: The Unintended Consequences of Consumer Research," In Frank R. Kardes, Paul M. Herr and J. Nantel (Eds.), *Applying Social Cognition to Consumer-Focused Strategy* (pp. 81-95). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Morrongiello, Barbara A., and Landa Mark (2008), "'Practice What You Preach': Induced Hypocrisy as an Intervention Strategy to Reduce Children's Intentions to Risk Take on Playgrounds," *Journal of Pediatric Psychology*, 33(10), 1117-1128.

Morwitz, Vicki G. (2005), "The Effect of Survey Measurement on Respondent Behaviour," *Applied Stochastic Models in Business and Industry*, 21(4-5), 451-455.

Morwitz, Vicki G., and Gavan J. Fitzsimons (2004), "The Mere Measurement Effect: Why Does Measuring Intentions Change Actual Behavior?" *Journal of Consumer Psychology*, 14(1&2), 64-74.

Morwitz, Vicki, G., Eric Johnson, and David Schmittlein (1993), "Does Measuring Intent Change Behavior?" *Journal of Consumer Research*, 20 (June), 46-61.

Obermiller, Carl and Eric R. Spangenberg (2000), "Improving Telephone Fundraising by Use of Self-Prophecy," *International Journal of Nonprofit and Voluntary Sector Marketing*, 5(4), 365-372.

Ofir, Chezy and Itamar Simonson (2001), "In search of negative customer feedback: The effect of expecting to evaluate on Satisfaction Evaluations," *Journal of Marketing*

Research 38 (May), 170-182.

Ofir, Chezy, Itamar Simonson and Song-Oh Yoon (2009), "The Robustness of the Effects of Consumers' Participation in Market Research: The Case of Service Quality Evaluations," *Journal of Marketing*, 73(6), 105-114.

Perkins, Andrew, Ronn J. Smith, David E. Sprott, Eric R. Spangenberg, and David C. Knuff (2008), "Understanding the Self-Prophecy Phenomenon," *Advances in Consumer Research – European Conference Proceedings*, 8, 462-467.

Sanbonmatsu, David M., Steven S. Posavac, Frank R. Kardes, and Susan P. Mantel (1998), "Selective Hypothesis Testing," *Psychonomic Bulletin & Review*, 5 (June), 197-220.

Sandberg, Tracy and Mark Conner (2009), "A Mere Measurement Effect For Anticipated Regret: Impacts on Cervical Screening Attendance," *British Journal of Social Psychology*, 48 (2), 221-236.

Schneider, Daniel, Alexander Tahk and Jon A. Krosnick (2007), "Reconsidering the Impact of Behavior Prediction Questions on Illegal Drug Use: The Importance of Using Proper Analytic Methods," *Social Influence*, 2(3), 178-196.

Schwarz, Norbert (1999), "Self-Reports: How the Questions Shape the Answers," *American Psychologist*, 54(2), 93-105.

Sherman, Steven J. (1980), "On the Self-Erasing Nature of Errors in Prediction," *Journal of Personality and Social Psychology*, 39 (August), 211-221.

Sherman, Steven J. (2008), "Should I Ask Our Children About Sex, Drugs, and Rock & Roll? A Different Conclusion," *Journal of Consumer Psychology*, 18(2), 96-101.

Simmons, Carolyn J., Barbara A. Bickart, and John G. Lynch, Jr. (1993), "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research*, 20 (September), 316-329.

Smith, Jennifer K., Alan S. Gerber and Anton Orlich (2003), "Self-Prophecy Effects and Voter Turnout: An Experimental Replication," *Political Psychology*, 24(3), 593-604.

Spangenberg, Eric R. (1997), "Increasing health club attendance through self-prophecy," *Marketing Letters*, 8(1), 23-31.

Spangenberg, Eric R. and Anthony G. Greenwald (2001). Self-prophecy as a behavior modification technique in the United States. In W. Wosinska, R. B. Cialdini, D. W. Barrett, & J. Reykowski (Eds.), *The practice of social influence in multiple cultures* (pp. 43-52). Mahwah, NJ: Erlbaum.

Spangenberg, Eric R. and Anthony G. Greenwald (1999), "Social Influence by Requesting Self-Prophecy," *Journal of Consumer Psychology*, 8 (1), 61-89.

Spangenberg, Eric R., Anthony G. Greenwald, and David E. Sprott (2008), "Will You Read This Article's Abstract? Theories of the Question-Behavior Effect," *Journal of Consumer Psychology*, 18 (2), 102-106.

Spangenberg, Eric R., and David E. Sprott (2006), "Self-Monitoring and Susceptibility to

the Influence of Self-Prophecy," *Journal of Consumer Research*, 32 (4), 550-556.

Spangenberg, Eric R., David E. Sprott, Bianca Grohmann and Ronn J. Smith (2003), "Mass Communicated Prediction Requests: Practical Application and a Cognitive Dissonance Explanation for Self Prophecy," *Journal of Marketing*, 67 (July), 47-62.

Sprott, David E., Ronn J. Smith, Eric R. Spangenberg, and Timothy S. Freson (2004), "Specificity of Prediction Requests: Evidence for the Differential Effects of Self-Prophecy on Commitment to a Health Assessment," *Journal of Applied Social Psychology*, 34 (June), 1176–90.

Sprott, David E., Eric R Spangenberg, and Andrew W. Perkins (1999), "Two More Self-Prophecy Experiments," in Eric J. Arnould and Linda Scott Price (Eds.) *Advances in Consumer Research, Vol. 26* (pp. 621-626), Provo, UT: Association for Consumer Research.

Sprott, David E., Eric R. Spangenberg, Lauren G. Block, Gavan J. Fitzsimons, Vicki G. Morwitz, and Patti Williams (2006), "The Question-Behavior Effect: What We Know and Where We Go From Here," *Social Influence*, 1 (2), 128-137.

Sprott, David E., Eric R. Spangenberg, and Robert J. Fisher (2003), "The Importance of Normative Beliefs to the Self-Prophecy Effect," *Journal of Applied Psychology*, 88 (June), 423–31.

Sprott, David E., Eric R. Spangenberg, David C. Knuff, and Berna Devezer (2006), "Self-Prediction and Patient Health: Influencing Health-Related Behaviors Through Self-Prophecy," *Medical Science Monitor*, 12(5), RA85-91.

Stutzer, Alois, Lorenz Goette and Michael Zehnder (2007), "Active Decisions and Pro-Social Behavior," *Working Paper 07-13*, Federal Reserve Bank of Boston.

Van Kerckhove Anneleen, Maggie Geuens, and Iris Vermeir (2009), "The Effects of Sequential Self-Predictions on Behavior Change," *Working paper*, Universiteit Gent.

Weiner, Bernard (1985), "'Spontaneous' Causal Thinking," *Psychological Bulletin*, 97 (1), 74-84.

Williams, Patti, Lauren G. Block, and Gavan J. Fitzsimons (2006), "Simply Asking Questions about Health Behaviors Increases both Healthy and Unhealthy Behaviors" *Social Influence*, 1 (2), 117-127.

Williams, Patti, Gavan J. Fitzsimons, and Lauren G. Block (2004), "When Consumers Do Not Recognize 'Benign' Intention Questions as Persuasive Attempts," *Journal of Consumer Research*, 31 (December), 540-550.

Wills, Thomas A. (1981), "Downward Comparison Principles in Social Psychology," *Psychological Bulletin*, 90(2), 245-271.

Wyer, Robert S. and Thomas K. Srull (1989), *Memory and Cognition in Its Social Context*. Hillsdale, NJ: Erlbaum.