

Does Measuring Intent Change Behavior?

VICKI G. MORWITZ
ERIC JOHNSON
DAVID SCHMITTLEIN*

Past research has established that, while self-reports of purchase intentions can predict behavior, various factors affect the strength of the intentions-behavior link. This article explores one such factor: the impact of merely measuring intent. Our specific question concerns the impact of measuring intent on subsequent purchase behavior. Prior research suggests a mere-measurement hypothesis: that merely measuring intent will increase subsequent purchase behavior. We also suggest a polarization hypothesis: that repeated intent questions will have a polarizing effect on behavior. The results reveal that the effect of merely asking intent to buy once is an increase in the subsequent purchase rate. The effect of repeatedly asking intent for those with low levels of intent is a decreased propensity to buy with repeated measurements. These two effects are reduced given prior experience with the product. The implications of these findings and opportunities for future research are discussed.

The road from reported intentions to action is rocky. As usually modeled in the behavioral sciences, statements of intended action often do not come to fruition. Situational influences and normative constraints are just two of the possible detours that prevent stated intentions from becoming reality. A large theoretical and empirical literature examines the relationship between statements of intent and subsequent behavior (Ajzen 1985; Fishbein and Ajzen 1975; Shepard, Hartwick, and Warshaw 1988). A newer stream of research suggests, however, that measuring intentions can have an influence of its own on subsequent behavior. One analogy is to the von Heisenberg uncertainty principle in physics, which states that by merely measuring the location of a particle, the position of the particle may change. By analogy, this paper explores the hypothesis that merely measuring someone's intent to purchase can change his or her likelihood of purchase.

As a starting point, we adopt the emerging view that values, intentions, and attitudes do not always exist in memory waiting to be assessed but may be constructed on demand, in response to a query, either internally generated by the consumer or externally, for example, by a market research questionnaire. This perspective

arises from both recent theoretical reviews in attitude theory (Feldman and Lynch 1988) and a number of studies in behavioral decision research (Fischhoff 1991; Fischhoff, Slovic, and Lichtenstein 1980; Kahneman and Snell 1990; Payne, Bettman, and Johnson 1992).

A key implication of this perspective is that answering a question has at least two effects. First, answering an attitude question can make the attitudes much more accessible and make resulting behavior more consistent with the underlying attitude. Second, answering an intention question not only increases attitude accessibility but can cause the respondent to engage in substantial cognitive work that may produce a subsequent change in attitudes and intentions. This processing in turn has been shown to change subsequent behavior (Hirt and Sherman 1985; Sherman 1980). Thus, there is evidence that answering an intent question can change behavior in two ways: by making attitudes more accessible and by changing the attitude itself.

Together, these two related lines of argument suggest that simply asking an intent question may change behavior. The objectives of our research, therefore, are (1) to examine the impact of measuring intent on subsequent behavior by comparing the purchase behavior of a group in which the members had not been asked for their intent to a comparable group for which an intent measure was obtained and (2) to investigate whether and how the repeated asking of intent with the same respondent, as is commonly performed by marketing research firms, affects purchase behavior.

In particular, this research will demonstrate that asking the intent question does alter behavior, and that consumers (*a*) purchase differently if asked any intent

*Vicki G. Morwitz is assistant professor of marketing at the Stern School of Business, New York University, New York, NY 10012. Eric Johnson and David Schmittlein are professors of marketing at the Wharton School, University of Pennsylvania, Philadelphia, PA 19104. The authors would like to thank Deborah Mitchell, John Lynch, and the three reviewers for helpful comments, and David Sullivan, formerly at Prodigy Services Company, Jerry Lieberman at General Motors, and Douglas Scott at Allison Fisher for providing the panel data analyzed.

questions, (b) differ in their purchase behavior depending on the number of times their intent has previously been measured, and (c) are less affected by intent measurement if they have previous product experience. The magnitudes and statistical significance of these effects are assessed with longitudinal panel data. The results suggest caution in both interpreting and using intentions to buy. They also have implications for the common practice of projecting sales levels from a sample survey of intentions to the relevant population.¹ Finally, a central goal of this work is to examine, in field study settings, the application of a constructive preference perspective to consumer choice for important big-ticket goods.

The next section reviews a broad set of empirical results that support this reasoning and develops a set of hypotheses. We then present the analysis of two large-scale panels that track intentions and purchases of consumers for two different important durables. Since the data are historical and the analysis is correlational, we also discuss the possible confounds and present analyses that attempt to control for them. Finally, we summarize the implications of these results for a constructive view of preferences, as well as their implications for practitioners, and close by outlining opportunities for further research.

MEASURING INTENT AND THE CONSTRUCTION OF PREFERENCES

How can asking an intent question change someone's behavior? An important consequence of answering an attitude question about a product is that the attitude itself becomes more accessible, as measured, for example, by reaction times (Fazio, Powell, and Williams 1989). Repeating an attitude question also produces an increase in accessibility (Kardes, Allen, and Pontes 1993). By analogy, we suspect that answering purchase intent questions might make the attitude underlying the intent more accessible. Since attitude accessibility is an important influence on attitude behavior consistency, we would expect intent questions to change behavior as well (Fazio 1989).

A related, but perhaps more relevant literature concerns people's predictions of their own future behavior. Sherman (1980) and his colleagues (Sherman et al. 1983) have shown that the very act of making a prediction about one's behavior can induce substantial be-

havioral change, when the subjects who made the prediction are compared to a control group that did not make a prediction. For example, subjects were asked whether they would volunteer three hours for a cancer society's fund-raising effort. Subjects who were asked to make predictions typically overestimated their willingness relative to the actual behavior of a group not asked to make a prediction. However, these errors are "self-erasing" in the sense that the group that had predicted high levels of behavior were indeed more likely to volunteer. Similarly, Greenwald et al. (1987) demonstrated that simply asking people whether they planned to vote increased their participation in actual elections. This leads to our first mere-measurement hypothesis:

H1: Merely measuring intent once will change the future purchases of a product when the group in which intent is measured is compared to a control group that has not been asked intent questions.

This mere-measurement hypothesis may seem a bit unlikely at first, given its application to a marketing research context. Unlike most of the behaviors that have been employed in social-psychological studies, such as spending three hours fund-raising or singing "The Star-spangled Banner" over the telephone, the purchase decisions that we will study are relatively important for the respondent, namely, purchasing a home computer or new automobile. In addition, the induction of intent is relatively weak: a single question about purchase intent embedded in an omnibus survey of many questions on a variety of topics. Yet, given the pervasiveness of these findings in social psychology, the hypothesis seems worth testing.

The direction of this effect will depend on two specific characteristics of the product and population. First, we need to know the nature of the additional cognitive work that will be invoked by the intention question. In many of the social-psychological studies examining intent and subsequent behavior, the activity itself is positively valued (voting, charitable donations, purchase of cable television services) and likely to invoke positive thoughts on a request for further consideration. For example, Gregory, Cialdini, and Carpenter (1982) demonstrate that consumers who merely considered a scenario in which they were cable television subscribers actually increased the rate at which they subscribed to the service.

We suspect that for many products this mental dialogue will generate positive thoughts. In fact Hoch (1984) used a similar intent question that both provoked primarily positive reasons and inhibited recall of negative reasons. Similarly, Kalwani and Silk (1982) suggest that the kinds of mental simulation (Kahneman and Tversky 1982) used by consumers might mediate the link between stated intent and subsequent purchase of durables. If increased thought provoked by an intention

¹Purchase intentions are routinely used for such purposes as forecasting (Infosino 1986), tracking (McNeil 1974), advertising research (MacKenzie, Lutz, and Belch 1986), new product testing (Silk and Urban 1978), and market segmentation (Sewall 1978). Many large market research firms such as Market Facts, NFO, and HTI repeatedly obtain intent to buy measures for a large range of products from households in their consumer mail panels (Lehmann 1989, pp. 255-286).

question is primarily positive, we suspect intentions could result in increased purchasing.

The second factor concerns the distribution in the population of prior attitudes. Because increased accessibility can increase the correspondence between attitudes and behavior, the aggregate effects of intention questions depend on whether the predominant attitude in a population is positive or negative. If it is positive, intention questions could increase purchases; if it is negative, they could decrease purchases.

Together, these two different processes suggest that intention questions could conceivably produce aggregate increases or decreases in sales. For most products where the thoughts induced by simulating ownership are not onerous and for which most people do not hold negative attitudes, however, we would expect increases to result from purchase intent questions. The two product classes we consider here, cars and personal computers (PCs), seem to have these aggregate characteristics.

Of course, at the individual level initial attitudes may not all be positive, and the effect of intent questions may well differ among individuals. Imagine someone who sees new car purchases as a wasteful use of the earth's resources or PCs as threats to privacy and autonomy. If they are asked an intent question, we might expect the attitude of such persons to become more negative and purchase likelihood to decrease. Thus, we offer a polarization hypothesis that the effect of intent questions will depend, at the individual level, on the initial level of intent:

- H2:** The impact of repeated measures of intent over time depends on the initial level of intent. (a) Those with high levels of intent will have increased purchase rates, and (b) those with low levels of intent will have decreased purchase rates.

This hypothesis is consistent with similar observations concerning attitudes, particularly Tesser's (1978) on the mere-thought hypothesis, which suggests that simply asking an attitude question polarizes attitudes. He suggests that "(a) for various stimulus domains persons have naive theories or *schemas* which make some attributes of the stimuli salient and provide rules for inferences regarding other attributes; (b) thought, under the direction of a schema, produces changes in beliefs, and these changes are often in the direction of greater schematic and evaluative consistency; (c) attitudes are a function of one's beliefs. Since thought tends to make beliefs more evaluatively consistent and attitudes are a function of beliefs, thought will tend to polarize attitudes" (p. 290). Similarly, Kardes et al. (1993) show that repeated attitude questions create increases in accessibility, and Downing, Judd, and Brauer (1992) have demonstrated that "attitudes that are expressed more frequently can become more extreme or polarized as a result" (p. 27).

Finally, we might expect these effects to depend on prior experience or knowledge of the domain. On the one hand, well-articulated attitudes (possessed generally by more experienced/knowledgeable persons) may show less susceptibility to attitude questions. (Feldman and Lynch [1988] make similar arguments.) On the other hand, research studies have indicated that attitude polarization as a function of thought is less likely when the attitude object is one for which people have a less developed schema (Tesser 1978, pp. 302-303). Presumably, more experienced or knowledgeable consumers have better-developed schemas and, *ceteris paribus*, will exhibit more attitude polarization. The *ceteris paribus* qualifier is important here, however. In the studies cited by Tesser, subjects appear not to have had preexisting, well-established attitudes toward the objects of interest. In short, we believe that Tesser's effect of schema will be overshadowed by the typical strength of prevailing attitudes, where such attitudes exist. Thus, we expect experience and knowledge to moderate Hypotheses 1 and 2 in the following way:

- H3:** The mere-measurement and polarization effects should be stronger among those consumers with less experience in the product class.

Similar results exist for predictions in general. For example, Hirt and Sherman (1985) asked both novices and knowledgeable fans to explain hypothetical outcomes of college football games prior to their occurrence. These explanations affect the later predictions of novices but not of experts. Similarly, Fazio and Zanna (1981) argue that when an attitude is based on direct product experience there will be higher correlations between attitude and behavior.

These three hypotheses have both theoretical and practical consequences. First, they help broaden our experience with the processes of preference formation and product choice. The view that product preferences are sometimes constructed is still in a relatively early stage of development, and the work here helps extend the range of application of those ideas. To the extent that the hypotheses are supported below we also provide empirical support for that general constructivist perspective. Our hypotheses have very practical consequences for the conduct of consumer surveys. That is, any intention-based purchase forecasts (even those validated intertemporally for the surveyed group) must be adjusted in predicting behavior for the (nonsurveyed) population at large. Further, the hypotheses point to the importance of survey and questioning methods that may help to reduce such intention-measurement effects. We will return to these points in the conclusion. In the next section the design used to test these hypotheses is described.

STUDY DESIGN

Data from two quasi experiments were used to test the hypotheses posited above. Rather than conduct a

formal experiment we used two multiwave panel survey data sets provided by sponsors of this research. These panel data reflect the results of a naturally occurring or quasi experiment about intentions to buy a durable good and subsequent acquisition of that good. Because of panel dynamics (members entering and exiting a panel over time), panel members varied in the number of times their intentions were measured. By monitoring their subsequent purchase patterns we could examine whether merely measuring intent affects purchase behavior.

In an analogous formal experiment one would randomly assign subjects to treatments; subjects would be assigned to receive either zero, one, or multiple intention measurements. The dependent measure, subsequent purchase of the product, would be monitored for subjects in all treatment groups. In this quasi experiment, a new sample of the population is recruited to replace those lost to attrition.

A primary goal of our research was to obtain results that would generalize from the particular findings of this study. To that end, we examined multiple products, time periods, and methods for assuring that treatment groups were equal in all respects other than the treatment itself. In particular, we analyzed data concerning intentions to buy an automobile, an established durable good purchased by many U.S. households, and a home PC, a relatively new durable good purchased by a minority of U.S. households. In both cases we measured timed intent (*intent* to buy in zero to six months, seven to 12 months, . . . , or never). The intention refers to purchase in the product category, as opposed to a particular brand. We examined the relationship between intent and behavior for three different six-month time frames for one of the two products (the automobile) in order to ensure that our results were not a chance occurrence in one particular time period. Finally, we used three alternative methods for ensuring that effects attributed to the number of intent measurements received did not stem from other differences across households. That is, we analyzed data weighted by two alternative criteria (as well as the unweighted data) and also report an analysis with a battery of household characteristics used as covariates.

The relatively large sample size here could not realistically be matched in any true experiment. (More than 40,000 households were used in analyzing each of our two product categories.) Further, realistic following of purchase behavior for durables would require an experiment lasting months at a minimum and probably years (as does the current study). Finally, the panel assembled here is designed to be representative of U.S. households, rather than a convenience sample from a particular age, geographic, or socioeconomic group. In summary, we believe that it would be most difficult to obtain results analogous to ours in a true experimental setting. Of course, as in any longitudinal quasi experiment, the results are open to alternative explanation if

households that elect to remain in a panel over time differ from those that drop out in ways that we have not controlled for above, and if such factors also significantly affect purchase behavior.

Data Description

For both products, data were obtained from seven waves of mail surveys sent to large consumer mail panels. The PC survey waves took place approximately every six months from July 1986 (wave 1) to January 1989 (wave 7).² The automobile surveys took place quarterly between the fourth quarter of 1987 and the second quarter of 1989. During each wave the panel households were asked their timed intent to buy an automobile or PC in the future. Specifically, the intent question in the PC survey during waves 4–7 reads:³

Do you or does anyone in your household plan to acquire a (another) personal computer in the future for use at home?

- Yes, in the next 6 months _____
- Yes, in 7 to 12 months _____
- Yes, in 13 to 24 months _____
- Yes, sometime, but not within
24 months _____
- No, but have considered
acquiring one _____
- No, will not acquire one _____

The intent question used in the automobile survey during waves 1–7 read:

When will the next new (not used) CAR (not truck or van) be purchased by someone in your household?

- 6 months or less _____
- 7–12 months _____
- 13–24 months _____
- 25–36 months _____
- Over 36 months _____
- Never _____

Both surveys also asked whether households owned a PC or had recently purchased an automobile. In addition, the PC survey asked panel households whether they had previously owned a home PC, whether any member of the household had ever used a PC at work or at school, and, if they did purchase, the month and year of purchase.

For many panel households the answers to the above intent questions were obtained during all seven waves. Because of attrition in the panels new replacement

²Wave 3 (July 1986) and wave 4 (July 1987) were one year apart.

³A rephrased version of this question was used during previous waves.

TABLE 1

NUMBER OF INTENT MEASURES RECEIVED IN THE PC SURVEY

Number of intent measures	Waves	Number of households
0	7	5,634
1	6, 7	2,138
2	5, 6, 7	9,512
3	4, 5, 6, 7	4,637
4	3, 4, 5, 6, 7	4,581
5	2, 3, 4, 5, 6, 7	15,113
6	1, 2, 3, 4, 5, 6, 7	4,928

households entered the panels during each wave. Therefore, some panel households answered the survey questions from only one wave, while others answered for two, three, or up to six survey waves.

The focus of this research is on the percentage of panel households that purchased during the six-month period following wave 6 (e.g., between July 1988 and January 1989 for the PC data). For the automobile data, we also analyze sales in the six-month periods following wave 4 (the third quarter of 1988) and wave 2 (the first quarter of 1988). All automobile purchases were considered (first and repeat). For the PC data we considered only first purchases. Because of the survey design we were able to identify purchases only for first-time PC buyers.⁴ Accordingly, households selected for analysis from the PC surveys were those that (1) responded to wave 7, (2) had not previously owned a PC at home, and (3) if they responded to wave 6 were not PC owners at that time.

The selected households were categorized by the number of times their purchase intent was measured. For instance, with the PC data the number of times intent was measured prior to wave 7, the survey waves the household responded to, and the number of panel households in each category are shown in Table 1.

Because we were not interested in the impact of specific numbers of intent measurements (e.g., two vs. three vs. four, etc.) we pooled households who received two or more intent questions into one group. Households in the PC study were further categorized by whether they had previous experience with the product at work and/or school. The sample sizes for the automobile data for the three time frames and for the PC data for households with and without previous product experience are reported in Exhibit 1.

By comparing the percentage of households that both acquired the product during the six-month interval of interest (e.g., between waves 6 and 7 for the PC data)

⁴We assumed a PC purchase had taken place when the household switched from being a nonowner to an owner from one survey wave to the next. It would not be possible to detect the purchase of an additional PC given the available information.

EXHIBIT 1

SAMPLE SIZES AND TIME PERIODS AVAILABLE FOR ANALYSIS

Automobile data (three different time periods were analyzed):	
Intention measured 1st quarter 1989; purchases monitored in 1st and 2d quarters of 1989:	
4,776 households never asked intent ^a	
3,571 households asked once	
45,023 households asked more than once	
Intention measured 3d quarter 1988; purchases monitored in 3d and 4th quarters of 1988:	
2,846 households never asked intent	
2,630 households asked once	
35,246 households asked more than once	
Intention measured 1st quarter 1988; purchases monitored in 1st and 2d quarters of 1988:	
3,605 households never asked intent	
8,254 households asked once	
23,387 households asked more than once	
Personal computer data (one time period was analyzed):	
Intention measured in July 1988; purchases monitored through January 1989:	
5,634 households never asked intent	
2,138 households asked once	
38,771 households asked more than once	
Intention measured separately for levels of product experience at work and/or school:	
Households without previous product experience:	
2,271 households never asked intent	
1,032 households asked once	
24,941 households asked more than once	
Households with previous product experience:	
3,363 households never asked intent	
1,106 households asked once	
13,830 households asked more than once	

NOTE.—In both data sets, the raw data and the data weighted by two different methods were analyzed.

^a These households did not have their intentions measured at the beginning of this period. Their intentions might have been measured once halfway into the purchase period. Since intention information is available quarterly, but sales information is only available bi-annually, we consider the number of times intentions were measured at the beginning of the six-month purchase period.

and were never asked intent to the corresponding percentage for households that were asked intent once, we could test Hypothesis 1, the mere-measurement hypothesis. (That is, we could observe the impact of asking the intent question during wave 6.) A test of multiple proportions was used to determine whether households whose intentions were measured once had a significantly greater purchase rate than households whose intentions were never measured. Households whose intent was never measured were not yet in the panel during wave 6, and so their wave 6 purchase intentions were not measured. Households that were asked once had just joined the panel in time for wave 6, and so their purchase intentions were available from this wave. In order for the comparison to be valid, these two groups must be similar on all other aspects. Statistical attempts to control for this are described below. To test Hypothesis 3, we repeated the above procedure separately for households with previous product experience and for those without such experience.

DOES MEASURING INTENT CHANGE BEHAVIOR?

Similarly, we compare the purchase rates of households with low or high initial intent receiving one or multiple intention measures, to assess the impact of repeatedly asking purchase intent questions over time. Logit models to predict purchase given the level of initial intent, the number of intent measures received, and their interaction were used to test the polarization hypotheses, Hypotheses 2a and 2b. To test Hypothesis 3 (the moderating effect of product experience) this procedure was repeated separately for households having and not having previous product experience.

In addition to the survey questions described above, several household demographic descriptors were available. The demographic variables used in this research are reported in Exhibit 2.

Since the data are correlational and not experimental, any results found might be due to demographic differences in groups receiving different numbers of intent measures. As will be described in the next section, the data were weighted on the basis of demographics to control for these differences.

Methods for Controlling Confounding Effects

Since the data are historical and correlational it is important to account for any possible confounding effects. For example, households that stayed in the panel for all seven survey waves might be systematically different from households that entered the panel later and hence were in the panel for a shorter amount of time. Ehrenberg (1960) demonstrated that length of time in a panel did not systematically affect the reported results on total purchases or brand shares for diary panel members. However, Schmittlein and Helsen (1988) showed that the propensity to leave a panel decreases with income, increases with male head of household's education, and is highest for professional workers and lowest for those retired or unemployed. (See Sudman and Ferber [1979, pp. 31-38] for a general discussion of bias in panels.) Therefore, households that remain in the panel for many waves, compared to other panel households, may have a higher proportion of households with heads with lower education and who are retired or unemployed. These households may be less likely to purchase an expensive durable good. In order to check for systematic differences between groups receiving different numbers of intent measurements, cross-tabulations were performed associating each of the household demographic variables (listed in Exhibit 2) with the number of times intent was measured. Systematic differences across the groups did exist. For both products, the groups who received more survey waves had an increased tendency to have households with older heads, retired heads, and widowed female heads.

EXHIBIT 2

HOUSEHOLD DEMOGRAPHIC INFORMATION USED IN ANALYSES

Personal computer data:
Presence/age of children
Age of female head of household
Age of head of household
Number of household members
Composition of household (e.g., male/female married, male alone)
Education of head of household
Number of hours head of household employed per week
Occupation of head of household
Type of housing
Total household income
Life cycle of household (e.g., new baby boomers, affluent elderly)
Marital status
Age of male head of household
Race
Census region
Socioeconomic status (e.g., female employed/male white collar)
Automobile data:
Size of household
Total household income
Age of head of household
Marital status
Life cycle of household (e.g., new baby boomers, affluent elderly)
Number of hours head of household employed per week
Education of head of household
Race
Occupation of head of household
Composition of household (e.g., male/female married, male alone)
Type of housing
Home ownership

Two methods were used to correct for this potential source of bias.⁵ (1) The data were weighted to be balanced on the life cycle variable. This was the variable with the largest chi-square statistic in the variable-by-number-of-intent-measures tabulations for both products. (2) Alternatively, the data were weighted to be balanced on age of the head of household. Recall that age is one of the most important discriminators between households that remain in the panel and households that elect to drop out. For both methods the data were balanced to match the distributions on the demographic variable in the cross section of households that responded to wave 7.⁶ In addition, the unweighted data were also analyzed, and the results for all the approaches are reported below.

⁵A third method for potential bias was also used on the PC data. Those households meeting a specific demographic profile were analyzed separately. The same qualitative pattern of results found with the other two methods for correcting potential bias was also found with this method.

⁶For example, suppose 10.5 percent of all the wave 7 automobile respondents were younger than 30 years old, while 15.0 percent of those whose intentions were never measured prior to wave 7 were under 30. The weight for respondents in this cell (those under 30 whose intentions were never measured) would be .700 (.105/.150). For a more detailed discussion on the methodology of weighting data, see Lehmann (1989, pp. 426-429).

An alternative to weighting the data is to include dummy variables for the demographic descriptors as covariates in logit models. This approach controls for individual differences in the demographic variables but requires the estimation of a greater number of parameters. We defer our discussion of this approach until later and begin our discussion using the more straightforward approach of weighting the data as described above.

EMPIRICAL RESULTS

In this section we report our empirical findings. Each section below corresponds to the results for the test of one hypothesis.

Hypothesis 1: The Mere-Measurement Hypothesis

We hypothesized that the impact of asking intent would be an increase in the purchase rate. Our argument was that consumers would overpredict purchases of these two products and, because of the self-erasing nature of these errors, purchase more. This hypothesis follows from arguments of increased polarization of attitudes, use of the simulation heuristic, and increased attitude accessibility, all caused by the mere measurement of intent.

For the automobile data, the purchase rates were obtained by simply noting for all households the percentage that reported purchasing a car in the six-month interval of interest. In the PC surveys a reported change in current ownership (from nonowner to owner) from one survey wave to the next was assumed to correspond to a purchase. Hence, for households whose intent was measured once these percentages were obtained by calculating the proportion of wave 6 (July 1988) non-PC households that became owners for the first time in wave 7. Households whose intent was never measured and that were reported owners in wave 7 were sent a follow-up survey from which the month and year of their PC purchase were obtained. The proportion of PC-owning households that purchased their first PC after July 1988 was obtained from the follow-up survey. This proportion was applied to the owners whose intent was never measured in wave 6 to estimate the overall purchase rate for those whose intentions were never measured.

Exhibit 3 reports the purchase rates both for households whose intentions were not measured and for those with intentions measured once. The results are reported for one time period and one weighting scheme for both products. Similar results were obtained for the other time periods and weighting schemes. In all cases the patterns are similar. Households that were not asked intent had a lower purchase rate than those that were asked once. This is consistent with Hypothesis 1 (and with Sherman's [1980] results). For example, Exhibit 3 shows that 2.4 percent of households not asked their

EXHIBIT 3
THE IMPACT OF MERE MEASUREMENT

Number of times intent was measured	% Acquiring the product	N
Automobile data: ^a		
0	2.40	4,776
1	<u>3.30</u>	<u>3,518</u>
Difference	.90	
	($p = .0069$)	
Personal computer data: ^b		
0	3.80	5,641
1	<u>4.48</u>	<u>2,139</u>
Difference	.68	
	($p = .0735$)	

^a Weighted by life cycle. Intent was measured in the first quarter of 1989; purchases were made the first half of 1989.

^b Weighted by life cycle. Intent was measured in July 1988; purchase was made by January 1989.

automobile intent bought in the next six months, while 3.3 percent of those whose automobile intent was measured once bought (a 37 percent increase). A test of multiple proportions shows that this observed increase is statistically significant ($p = .0069$). For PCs, asking intent once increased the purchase rate from 3.80 percent to 4.48 percent, an 18 percent increase with p -value of .0735.

An increased purchase rate with a single intent measurement was observed for all nine time period and weighting scheme combinations for the automobile data and all three weighting schemes for the PC data. Thus, we clearly have directional support for Hypothesis 1. For the automobile data the observed increases were significant ($p < .05$) for six of nine time period and weighting scheme combinations. Pooled across time periods, the observed increases were significant ($p < .01$) for all three weighting schemes. For the PC data the observed increases for the three weighting schemes were not significant at the .05 level ($p = .456$; $p = .0735$; $p = .1075$). These results as well as the results for our other hypotheses are summarized in Exhibit 4. Our results support Hypothesis 1. The proportion of households that acquired a product was greater for households whose intentions were measured once than for those whose intentions were never measured.

Hypothesis 2: Does Repeated Measurement Polarize Purchases?

We hypothesized that repeated intent measurement would polarize subsequent purchase rates. That is, it would increase purchase rates for households with high initial intent and decrease purchase rates for households with low initial intent. In order to test each of Hypotheses 2a and 2b a logit model was estimated to predict purchase given the level of initial intent and the number of intent measures to which the household had

EXHIBIT 4
SUMMARY OF IMPACT OF MERELY MEASURING INTENT

	Directional support ^a	Significance ^{a,b}	Pooled significance across time ^b
Hypothesis 1: mere measurement:			
Automobile	9 of 9	6 of 9	3 of 3
Personal computer	3 of 3	0 of 3	
Hypothesis 2a: polarization hypothesis for households with high initial intent:			
Automobile	6 of 9	3 of 9	1 of 3
Personal computer	3 of 3	0 of 3	...
Hypothesis 2b: polarization hypothesis for households with low initial intent:			
Automobile	9 of 9	6 of 9	3 of 3
Personal computer	3 of 3	2 of 3	...
Hypothesis 3 for Hypothesis 1: is mere measurement mediated by product experience?			
Automobile
Personal computer	3 of 3	3 of 3	...
Hypothesis 3 for Hypothesis 2a: is polarization effect Hypothesis 2a mediated by previous product experience?			
Automobile
Personal computer	3 of 3	3 of 3 ^c	...
Hypothesis 3 for Hypothesis 2b: is polarization effect Hypothesis 2b mediated by previous product experience?			
Automobile
Personal computer	3 of 3	3 of 3 ^c	...

NOTE.—Entries represent the number of positive instances out of the total number of time periods and weighting schemes.

^a Across time and weighting schemes for automobiles; across weighting schemes for PCs.

^b $\alpha = .05$.

^c Examination of significance levels in a separate task from directional support task.

been exposed. To uncover the main and interaction effects of each (polytomous) variable, the following coding scheme was used. Two dummy variables represented initial intent—one for high initial intent (corresponding to intent to buy in the interval of interest, the next six months), and one for low initial intent (corresponding to intent never to buy). (The omitted or baseline level corresponds to intermediate intent—intent to buy beyond the immediate six-month time period.) Another dummy variable captured the number of intent measures (one or multiple). Finally, the model included interaction terms for initial intent by number of times measured. The sign and significance of the coefficients of these interaction terms were used to test our hypotheses.

Hypothesis 2a: Does Repeated Intent Measurement Increase Purchases for Households with High Initial Intent?

Hypothesis 2a stated that purchase rates would be higher with multiple than with single intent measurements for households with high initial intent. The results for both products for one time period and one weighting scheme are shown in Exhibit 5. For the automobile data, the purchase rate was 25.26 percent for

households whose intent was measured once and 28.65 percent for those whose intentions were measured more than once. Thus, the purchase rate increased by 13.4 percent with multiple intent measurements. The coefficient for the interaction of high initial intent and number of times measured was positive but not statistically significant ($p = .298$). Similarly, the purchase rate increased by 10.9 percent with multiple intent measures in the PC data (from 29.54 percent to 32.77 percent). As with the automobile data this increase was not significant ($p = .1860$).

The change in purchase rate for households with high initial intent with multiple intent measurements was positive for six of nine time periods and weighting schemes for the automobile data and all three weighting schemes for the PC data. Pooling the automobile data across time resulted in a positive change for two of three weighting schemes. The evidence does not provide overwhelming directional support for Hypothesis 2a. In addition, the changes were significant (in the hypothesized direction) for only three of nine time periods and weighting schemes for the automobile data. Pooled across time, these changes were only significant for one of the three weighting schemes. The changes were also not significant for any of the weighting schemes for the

EXHIBIT 5
DOES REPEATED MEASUREMENT POLARIZE RESPONSES?

Stated intent	No. of times intent was measured		Significance (<i>p</i>) of interaction
	Once ^a	Multiple ^b	
Automobile data:^c			
Intend to buy/0–6 months	25.26	28.65	.2980
Intend to buy/>6 months	4.64	4.81	
Never intend to buy	.49	.15	.0079
Overall	3.92	3.99	
Personal computer data:^d			
Intend to buy/0–6 months	29.54	32.77	.1860
Intend to buy/>6 months	4.98	4.32	
Never intend to buy	2.13	1.13	.0382
Overall	4.48	3.09	

^a For automobile data, *N* = 2,629; for PC data, *N* = 2,140.

^b For automobile data, *N* = 35,243; for PC data, *N* = 38,839.

^c Weighted by life cycle. Intent was measured in the third quarter of 1988; purchases were made over the second half of 1988.

^d Weighted by life cycle. Intent was measured in July 1988; purchase was made by January 1989.

PC data at the .05 level. These results are summarized in Exhibit 4.

Hypothesis 2a was not supported. We did not find sufficient evidence for an additional increase in sales with repeated intent measurement for households with high initial intent. This seems surprising, given the increases seen in Exhibit 5. There are, however, two quite plausible reasons that this result does not reach statistical significance. One possible reason for the lack of significance may be the relatively small sample of households with high initial intent. Although in general the sample sizes in this research are quite large, only about 4 percent of the automobile sample and 1 percent of the PC sample have high initial intent. Another potential cause of the lack of significance may be the correlation between prior experience with the product and having high initial intent. We will return to this when we discuss the moderating effect of prior product experience in Hypothesis 3.⁷

Hypothesis 2b: Does Repeated Intent Measurement Decrease Purchases for Households with Low Initial Intent?

Hypothesis 2b stated that purchase rates would be lower with multiple than with single intent measure-

⁷One might also wonder whether the nonsupport of Hypothesis 2a might be related to our operationalization of high intent vs. low intent through a timed intent question (e.g., intent to buy within six months vs. never), rather than through a more "classic" intent scale (intention to buy in the next year on a 0–10 scale). One of us (V.G.M.) has examined the predictive validity of timed intent questions (Morwitz 1993) and has found them to be indicators predominantly of individuals' (stationary) purchase rates for fixed time intervals, rather than of their purchase times per se.

ments for households with low initial intent. Low initial intent was defined as a stated intent never to buy the product. The results for both products for one time period and one weighting scheme are shown in Exhibit 5 (in the rows labeled "Never intend to buy"). For the automobile data, the purchase rate was 0.49 percent for households whose intent was measured once and 0.15 percent for households whose intentions were measured multiple times, a decrease of 69.4 percent. The coefficient for the interaction of low initial intent and number of times measured was both negative and statistically significant ($p = .0079$). Similar results were obtained with the PC data. The purchase rate decreased by 46.9 percent with multiple intent measures (from 2.13 percent to 1.13 percent). This decrease was also significant ($p = .0382$).

The change in purchase rate for households with low initial intent with multiple intent measurements was negative for all nine time periods and weighting schemes for the automobile data and all three weighting schemes for the PC data. Pooling the automobile data across time resulted in a negative change for all three weighting schemes. Thus, there is clear directional support for Hypothesis 2b. The changes were statistically significant for six of nine time periods and weighting schemes for the automobile data. Pooled across time, these changes were significant for all three weighting schemes. The changes were significant at the .05 level for two of three weighting schemes for the PC data ($p = .0041$; $p = .0367$; $p = .0812$). There was enough evidence to support Hypothesis 2b. A decrease in sales was observed to occur with repeated intent measurement for households with low initial intent.

Hypothesis 3 for Hypothesis 1: Does Prior Product Experience Moderate the Mere-Measurement Effect (Hypothesis 1)?

Previous experience with the product was hypothesized to reduce the effect in Hypothesis 1. That is, we expect that those who have not had product experience in the past will be more influenced by intent measurement so that the proportional increase in sales due to a single intent measurement will be greater for these households than for those that have had experience with the product.

Exhibit 6 shows the effect of a single intent measurement on PC purchases separately for households having and households not having previous product experience, for one of the three weighting schemes. Measuring intent once increased sales by 20.73 percent for households with past product experience and by 45.27 percent for households without past product experience. Thus, we have directional support for Hypothesis 3 as it pertains to Hypothesis 1; the mere-measurement effect is dampened by previous product experience. Furthermore, the probability value or significance level on the test of two proportions was higher (less significant) given

DOES MEASURING INTENT CHANGE BEHAVIOR?

EXHIBIT 6
IS MERE MEASUREMENT MODERATED
BY PRODUCT EXPERIENCE?

No. of times intent was measured	% Acquiring the product	N
Households with previous product experience:		
0	5.79	3,088
1	6.99	1,032
% Increase	20.73	
Significance of difference	$p = .0823$	
Households without previous product experience:		
0	1.48	2,603
1	2.15	1,108
% Increase	45.27	
Significance of difference	$p = .0735$	

NOTE.—Data are for PCs, weighted by life cycle. Intent was measured in July 1988; purchase was made by January 1989.

past experience.⁸ Exhibit 4 summarizes the findings across weighting schemes. The proportional increase in acquisitions was greater for households that had not had past PC experience for three of three weighting schemes. The probability values of the impact of asking once was smaller for those households that had not had past PC experience for two of three weighting schemes.⁹ Hypothesis 3 is supported as it relates to Hypothesis 1. The impact of measuring intent once was greater for households that did not have past product experience than for those with product experience.

Hypothesis 3 for Hypothesis 2: Is Polarization Moderated by Product Experience?

We hypothesized that previous experience with the product would dampen the effects hypothesized in Hypotheses 2a and 2b above. Specifically, we expected that, compared to households with past product experience, those who had never before used the product would have a higher proportional increase in sales with multiple intent measures if they had high initial intent and a greater proportional decrease in sales with multiple intent measures if they had low initial intent.

In order to test for these moderating effects, we computed the proportional changes in purchase rates from

⁸A probability value reflects both the magnitude in proportional differences and sample size. If the difference is held constant, increasing the sample size decreases the p -value. Since the sample size is greater for households with past product experience, our reported finding of decreased significance given past product experience is conservative.

⁹The one case where the p -value was not smaller with past PC experience was also the case where the difference in sample sizes was greatest, as discussed in n. 8. This sample size effect may thus account for the one contradictory finding.

one to multiple intent measurements separately for households with and without past product experience. As for the previous analysis of Hypothesis 3 for Hypothesis 1, we can do this only for the PC data set. We expected higher absolute proportional changes for households without previous product experience. Separate logit models were also fit for households having and households not having previous product experience. The models contained the same terms as described in section Hypothesis 2. We anticipate the probability values for the interaction terms to be higher (less significant) for households with past experience.

Hypothesis 3 for Hypothesis 2a: Is the Effect in Hypothesis 2a Dampened by Previous Product Experience?

Exhibit 7 shows the results for our test of Hypothesis 3 for Hypothesis 2a (along the rows labeled "Intend to buy/0-6 months") for the PC data for one weighting scheme. A 7.09 percent increase in the sales rate (from 34.86 percent with one intent measure to 37.33 percent with multiple measures) was observed for households with previous experience with a PC at work or school. For households without such product experience an 83.45 percent increase in the percentage purchasing was observed given multiple intent measures (from 11.42 with one intent measure to 20.95 with multiple intent measures). The proportional increase in sales for households with high initial intent was greater for households that had not had past PC experience for all three weighting schemes. These results support Hypothesis 3. The proportional increase in sales was dampened by previous product experience.

Recall that we did not find support for Hypothesis 2a based on the interaction terms in the logit model. Here, however, we expect greater significance for the interaction terms in the logit model for households without past product experience. The probability value for the interaction of high initial intent and number of times measured was .2714 for households with product experience and .1639 for households without product experience. Although neither term was significant at conventional levels, the interaction term was more significant for households without past PC experience.¹⁰ This was true for all three weighting schemes.

Hypothesis 3 as it relates to Hypothesis 2a is supported. The proportional increase in sales due to repeated intent measurement for households with high initial intent was greater for households that had not had previous product experience than for those that had had previous product experience. This result also explains the lack of support found for Hypothesis 2a. More than 75 percent of the households with high initial

¹⁰As before, because of differences in sample sizes this finding is conservative.

EXHIBIT 7
IS POLARIZATION MODERATED BY PRODUCT EXPERIENCE?

Stated intent	No. of times intent was measured (%)		Significance (p) of interaction	% Change, once vs. multiple
	Once ^a	Multiple ^b		
Households with previous product experience:				
Intend to buy/0–6 months	34.86	37.33	.2714	+7.09
Intend to buy/>6 months	6.65	5.96		
Never intend to buy	3.04	2.63	.4664	–13.49
Overall	6.99	5.63		
Households without previous product experience:				
Intend to buy/0–6 months	11.42	20.95	.1639	+83.45
Intend to buy/>6 months	2.40	2.36		
Never intend to buy	1.76	.61	.0092	–65.34
Overall	2.15	1.38		

NOTE.—Data are for PCs, weighted by life cycle. Intent was measured in 1988; purchases were made by January 1989.

^a For households with previous product experience, $N = 1,032$; for households without previous product experience, $N = 1,108$.

^b For households with previous product experience, $N = 15,630$; for households without previous product experience, $N = 23,209$.

intent had previous PC experience. Since the measurement effect is diminished with previous product experience and most households had that experience, it is logical that we did not see an overall increase in purchases as hypothesized in Hypothesis 2a.

Hypothesis 3 for Hypothesis 2b: Is the Effect in Hypothesis 2b Dampened by Previous Product Experience?

The rows labeled “Never intend to buy” in Exhibit 7 show the results of our test of Hypothesis 3 for Hypothesis 2b for the PC data with one weighting scheme. A 13.49 percent decrease in the sales percentage (from 3.04 with one intent measure to 2.63 with multiple measures) was observed for households having previous experience with a PC at work or school. For households without experience a 65.34 percent decrease was observed (from 1.76 with one intent measure to 0.61 with multiple intent measures). The proportionate decrease in sales for households with low initial intent was greater for households that had not had past PC experience for all three weighting schemes. These results support Hypothesis 3 as it relates to Hypothesis 2b. The proportionate decrease in sales was dampened by previous product experience.

The probability value for the interaction of low initial intent and number of times measured was .4664 for households with product experience and .0092 for households without product experience. The decrease was significant only for households without previous product experience. This was true for all three weighting schemes.

Hypothesis 3 for Hypothesis 2b is supported. The proportional decrease in sales due to repeated intent measurement among households with low initial intent

was greater when the household did not have previous product experience than among those with previous product experience. In short, our polarization effect is indeed greater for those without previous product experience.

Overview of Purchase Results

The results from the previous sections are summarized in Exhibit 4. Clearly the mere act of measuring intent does affect purchase behavior. Because we employed multiple products, time periods, and weighting schemes, we cannot dismiss the results as chance occurrences or as artifacts of a particular quasi experiment. The results provide clear overall weight for all but one of our hypotheses.

The analyses reveal that the effect of merely asking intent to buy once is an increase in the subsequent purchase rate. This effect was dampened by previous experience with the product. The effect of repeatedly asking intent to the same household was a decreased likelihood of purchase for stated nonintenders. Finally, the hypothesized effects of obtaining repeated intent measurements were greater for households without past product experience (a greater increase in sales for those with high initial intent and a greater decrease in sales for those with low initial intent).

We hypothesized that multiple intent measurements would lead to increased sales for households with high initial intent. Our results did not support this hypothesis. In retrospect, the following thoughts are noteworthy. First, our tests for Hypotheses 2a and 2b were conservative. Our logit models included terms for high initial intent and low initial intent. Therefore, the baseline for comparison is households with midlevel initial intent (intend to buy sometime, but later than the im-

DOES MEASURING INTENT CHANGE BEHAVIOR?

mediate six-month period of interest). Thus, in determining the significance of the interaction terms we are comparing both slopes to the horizontal line corresponding to the interaction of midlevel intent and number of times measured. Had we instead compared the two slopes to each other, our likelihood of finding significant interaction terms would have been greater.¹¹ Second, we demonstrated that households with previous product experience are less likely to be influenced by the mere act of expressing their intent than households without that experience. A respondent who states an intent to buy an expensive durable good within the next six months presumably is more knowledgeable about that product than someone stating a lower level of intent. Finally, more extensive mental simulation may occur for abnormal or atypical judgments (here, non-intentions) than for normal or typical judgments (e.g., intentions).¹² It seems normal for consumers to think about purchase intention judgments when there is some probability of purchase. In contrast, thinking about and forming nonintention judgments when there is a very low probability of purchase appears a bizarre exercise unlikely to be performed spontaneously. Such judgments are likely to be considered and formed only in response to explicit inquiries, for example, those embedded in a market research survey. So these non-intention judgments are only likely to be formed under relatively unusual, abnormal conditions, and these conditions may prompt greater levels of mental simulation and (as a result) judgmental polarization. This would lead, then, to the polarization of results for non-intenders (Hypothesis 2b) but not to such a result for intenders (Hypothesis 2a), again accounting for our empirical findings.

Experience with the product class is shown to moderate the hypothesized effects, and, in fact, consumers with low levels of experience are most likely to show both the mere-measurement and the polarization effects. This result is interesting in two different ways. First, since it replicates an effect that is found in many laboratory studies, it increases our confidence that the mechanism that produces the shift in purchases due to measurement of intent is the same as in the laboratory studies. In turn, it weakens alternative explanations because of the historical and correlational nature of our data. Second, this result demonstrates an important boundary condition on the influence of merely measuring intent, a point to which we will return later.

¹¹For example, for the weighting schemes and time periods in Exhibit 5, these interactions are indeed significant (for the PC data, $p = .0212$; for the automobile data, $p = .0058$). We still prefer our more conservative test of the hypothesized effect, so we do not pursue this avenue further. Note that, while comparing the slopes directly to each other would indeed suffice to demonstrate a general sort of *polarization* effect of the form hypothesized, it would not suffice to support the specific statements of that effect contained in Hypotheses 2a and 2b.

¹²This line of explanation was suggested to us by a reviewer.

The experience-effect findings also increase our confidence in our tests of the mere-measurement and the polarization hypotheses. Although we weighted the data to control for differences between groups, our study is correlational rather than experimental. A pathological example could most likely be constructed where groups receiving different numbers of intention measures also differ in some other important respect, with the latter causing the mere-measurement result. For example, if households that entered the panel in time for one wave were of a demographic group with a higher likelihood of buying an expensive durable good than households that did not receive an intention measure, then the demographic differences alone could account for a higher purchase among households receiving a single intention measure. However, since prior product experience was shown to dampen the effect, it is less likely that the observed differences in purchase rate could be due to demographic differences in the groups. It is harder to construct a scenario where demographic differences could account for the polarization results, and, given that experience was also shown to dampen the polarization results, we are confident that our findings are indeed the result of differences across households in the number of intention measures received and the level of initial intent and not due to systematic differences in other household characteristics. Although the experience effect did increase our confidence in all our results, we have also employed an additional approach to assure ourselves that the observed results were not caused by an unexamined confound of the sort discussed above. This methodology is described below.

CONTROLLING FOR INDIVIDUAL DIFFERENCES IN HOUSEHOLD DEMOGRAPHICS

In the previous sections we weighted the data on the basis of one demographic variable at a time in order to control for differences between the groups. While weighting on one variable also appeared to control for differences in the other demographic variables and facilitated very straightforward analyses of measurement effects, we also employed a methodology that controls simultaneously for multiple covariates.¹³ Our illustration here will use the automobile data. This approach incorporates the household demographic descriptors as covariates in a logit model for the purchase/no purchase outcome. These covariates control for individual differences in all of the demographic variables included in the models. The specific demographics included are the first nine variables for the automobile data in Exhibit 2, selected from the overall set in order to minimize collinearity. Dummy variables were used to code the nominal descriptors. This results in 62 demographic

¹³We would like to thank a reviewer for suggesting this approach.

EXHIBIT 8

A COMPARISON OF TWO DIFFERENT APPROACHES FOR CONTROLLING
 CONFOUNDING EFFECTS: WEIGHTING THE DATA (W)
 VERSUS INCLUDING COVARIATES IN A LOGIT MODEL (C)

	Directional support ^a		Significance ^b		Average β^a		Average p -value ^c	
	W	C	W	C	W	C	W	C
Hypothesis 1: mere measurement	9 of 9	3 of 3	6 of 9	2 of 3	NA	NA	.1308	.0556
Hypothesis 2a: polarization hypothesis for households with high initial intent	6 of 9	2 of 3	3 of 9	1 of 3	-.0411	+.1899	.1627	.1946
Hypothesis 2b: polarization hypothesis for households with low initial intent	9 of 9	3 of 3	6 of 9	1 of 3	-.6589	-.5438	.1239	.2050

NOTE.—Data are for automobiles. Entries for "directional support" and "significance" represent the number of positive instances out of the total number of time periods and weighting schemes.

^a Across three time periods and three weighting schemes when weighting the data; across three time periods for including covariates in a logit model.

^b $\alpha = .05$.

^c Averaged across the time periods and/or weighting schemes for which there was directional support for the hypotheses.

parameters to be estimated. The results of these analyses for the mere-measurement and the polarization hypotheses are described below.

Hypothesis 1: The Mere-Measurement Hypothesis

In order to test the mere-measurement hypothesis, we estimated a logit model predicting purchase/no purchase given number of intention measures (zero or one) and the 62 demographic covariates. A separate model was estimated for each of the three time periods. The mere-measurement hypothesis is supported if the coefficient for the number of intention measures is positive and statistically significant (according to a one-tailed test). The coefficients were indeed positive for all three time periods ($\beta = .1632$; $\beta = .2957$; $\beta = .2316$) thus providing directional support for our hypothesis. This result is consistent with our findings based on the weighted data, where all nine test statistics (based on a test of two proportions for three time periods and three weighting schemes) were in the hypothesized direction.

The logit model coefficient for the number of intention measures was statistically significant for two of three time periods ($p = .1152$; $p = .0261$; $p = .0256$), providing support for the mere-measurement hypothesis. This result is also consistent with the results of our test of two proportions based on the weighted data, where the test statistic was statistically significant in six of nine time periods and weighting schemes. The average probability value from the test of proportions for the weighted data was .1308 (averaged across the three time periods and three weighting schemes) while the average value based on the logit models was .0556 (averaged across the three time periods).

Our additional analysis provides greater confidence that the increased purchase rate associated with a single

intent measure did in fact result from the measurement and not from other systematic differences in the groups not accounted for by weighting the data on a single variable. In fact, after more carefully controlling for individual differences in demographics by employing a logit framework we find an even stronger mere-measurement effect (according to the lower average p -value for this approach). These findings are summarized in Exhibit 8.

Hypothesis 2: Does Repeated Measurement Polarize Purchases?

Earlier in the article we tested the polarization hypotheses by employing a logit model on the weighted data to predict purchase given number of intention measures (one or many), initial intent (high, middle, or low), and the interaction of number of measures by initial intent. For purposes of this discussion we will call this model the "weighted logit model." In order to more carefully control for the household descriptors we reestimated this logit model on the unweighted data but additionally included variables for the demographic descriptors as described in the previous section. We call this expanded model the "covariate logit model." The results for our polarization hypotheses are described below.

Hypothesis 2a: Does Repeated Intent Measurement Increase Purchase Rates for Households with High Initial Intent?

The hypothesis that repeated intent measurement increases purchase rates for households with high initial intent is supported if the coefficient for the interaction of number of measures and high initial intent is positive and statistically significant on the basis of a one-tailed

test. On estimating the covariate logit model we found this coefficient was positive for two of three time periods ($\beta = +.5051$, $\beta = +.0799$, $\beta = -.0153$). As with the weighted logit model, we again have some but not overwhelming support for Hypothesis 2a. The interaction coefficient was statistically significant for only one of three time periods under the covariate logit model. As with the weighted data approach, we do not find sufficient evidence to support Hypothesis 2a.

Hypothesis 2b: Does Repeated Intent Measurement Decrease Purchase Rates for Households with Low Initial Intent?

In this case, Hypothesis 2b is supported if the coefficient for the interaction of number of measures and low initial intent is negative and statistically significant on the basis of a one-tailed test. For the covariate logit model this coefficient was negative in all three time periods ($\beta = -.1681$; $\beta = -1.2052$; $\beta = -.2582$). These results provide directional support for Hypothesis 2b consistent with our previous weighted logit findings.

The interaction coefficient for number of intention measures by low initial intent was statistically significant for one of three time periods under the covariate logit model ($p = .4107$; $p = .0100$; $p = .1943$). Thus, the statistical confidence for Hypothesis 2b could be stronger, but it does not seem that the observed decrease in purchase rates with multiple intention measures can be attributed only to demographic differences in the groups. After controlling for demographic descriptors in a logit framework the absolute value of the average coefficient for the interaction term of interest was slightly greater ($\beta = -.6589$) than in the simpler weighted logit model ($\beta = -.5438$). The average probability level for the coefficient when it was in the hypothesized direction was lower (more significant) for the weighted logit model ($p = .1239$) than for the covariate logit model ($p = .2050$), presumably because of the increased number of parameters estimated in the latter. On the basis of the covariate logit model, we have some support for Hypothesis 2b.

Exhibit 8 summarizes the comparison between the two different approaches for controlling confounding effects. These results give us more confidence in the earlier finding based on weighting the data. One would expect the significance of our results to be decreased under the covariate logit models. We would have been concerned if the magnitude of the coefficients had also been lower when covariates were included in the covariate logit models than in the weighted logit models. Since the magnitude of the average coefficient actually increased in Hypothesis 2a and only slightly decreased in Hypothesis 2b, we conclude that measuring intent does indeed change behavior. The observed differences in purchase rates across groups do not seem attributable to differences in the demographic profiles of the groups.

CONCLUSIONS

The primary purpose of this research was to assess the impact of merely measuring intent. It is the practice of several market research firms to maintain large consumer mail panels and to routinely measure the panel households' intent to buy a variety of products. In accordance with past research, asking intent once (vs. not asking intent) was expected to increase the subsequent purchase rate. In this research such an increased purchase rate was observed. This finding is disturbing for those using intentions and subsequent purchase behavior to predict population (not sample) purchase patterns—the usual practical purpose for collecting intention data. If the effect of measuring intent truly is to increase purchasing, caution must be taken in interpreting stated intentions to buy. For example, intentions are often used as the dependent variable in new product testing, as a proxy for future purchase. Reliance on this measure can result in an overestimate of future sales if the consumers used in the test had a higher propensity to buy simply because their intent was measured. The effects observed could be exacerbated in a new product test, where most respondents would be unfamiliar with the product and therefore more influenced by intent measurement. Future research is needed to further examine the impact of measuring intent. An experiment could be designed to assess the magnitude and significance of this impact. Although several methods were used in this research to eliminate the demographic differences in the groups, systematic differences across the groups may still exist. An experiment would rule out the possibility of other confounding effects.

We stated earlier that the hypothesis that measuring intent would change behavior seemed to have a rather stringent test in these panels. Given both the importance of the purchase decision and the modest manipulation, we thought such effects, despite numerous laboratory and field experiment demonstrations, would be small. Given these prior beliefs, the size of these effects is surprising. In some analyses merely measuring intent increased purchasing rates by as much as 84 percent for no-experience PC shoppers who had some intention of buying. Stating a negative intent decreased purchases by as much as 69 percent for an even-bigger-ticket item, new car purchases.

All told, these results strongly support the evolving view that, even in important consumer decisions, values are sometimes constructed to fulfill the demands of the environment—in this case, questions asked by a market research firm. More important, our results indicate that such questions can have a long-term and sizable behavioral impact with important economic consequences. In a sense, these results relate to recent research on explanation and mental simulation. As Koehler has summarized (1991, p. 500), “Explaining is believing”; that is, persons asked to provide an argument supporting a hypothesis usually find the hypothesis more

plausible as a result. In our research, merely thinking (as opposed to explaining) may have an analogous effect, and the effect may follow through to behavior rather than stopping at beliefs or attitudes.

In a recent editorial, Lutz (1991) encouraged greater attention to "substantive phenomena identified as consumer behaviors" and stressed "the value of natural research settings." Our work in this article has stressed the description and understanding of a basic substantive consumer behavior pattern, as well as contributed to the conceptual work on attitude and preference formation. It has also brought together "behavioral" (largely cognitive psychological) perspectives with real consumers purchasing real products in a natural setting over an extended time period. Another example of a systematic, generalizable (large-scale) study of real consumer behavior based on behavioral perspectives can be found in Hardie, Johnson, and Fader (1992). It appears to us that there are great opportunities for such research, and we hope that Lutz's editorial encourages more of this kind of work.

Clearly, our results will raise more questions that are deserving of future research, and they would be bolstered by replication with other products and populations. Important methodological work remains concerning how market researchers should control for the possibility that measures of intent are reactive. We feel, however, that the most important remaining work concerns our understanding of the construction of values. The current research suggests the need to connect results from attitude research, research on memory, and behavioral decision research. Among the important questions that our current data cannot answer is the relationship between the cognitive process consumers go through when answering intent questions and any resulting change in the representation of that product and product class in memory. Research such as that conducted by Blair and Burton (1987), Hippler, Schwarz, and Sudman (1987), and Tourangeau and Rasinski (1988) would be needed to better understand these processes.

We hope that the demonstration of the effects in this article will engender fruitful cooperation between practitioners who employ intent measures and others who may find these effects to be interesting ways of studying the process of value construction.

[Received August 1991. Revised August 1992.]

REFERENCES

- Ajzen, Icek (1985), "From Intentions to Actions: A Theory of Planned Behavior," in *Action Control: From Cognition to Behavior*, ed. Julius Kuhl and Jurgen Beckmann, Berlin: Springer, 11-39.
- Blair, Edward and Scot Burton (1987), "Cognitive Processes Used by Survey Respondents to Answer Behavioral Frequency Questions," *Journal of Consumer Research*, 14 (September), 280-288.
- Downing, James W., Charles M. Judd, and Markus Brauer (1992), "Effects of Repeated Expressions on Attitude Extremity," *Journal of Personality and Social Psychology*, 63 (July), 17-29.
- Ehrenberg, Andrew S. C. (1960), "A Study of Some Potential Biases in the Operation of a Consumer Panel," *Applied Statistics*, 9 (March), 20-27.
- Fazio, Russell H. (1989), "On the Power and Function of Attitudes: The Role of Attitude Accessibility," in *Attitude Structure and Function*, ed. Anthony R. Pratkanis et al., Hillsdale, NJ: Erlbaum.
- , Martha C. Powell, and Carol J. Williams (1989), "The Role of Attitude Accessibility in the Attitude-to-Behavior Process," *Journal of Consumer Research*, 16 (December), 280-288.
- and Mark P. Zanna (1981), "Direct Experience and Attitude-Behavior Consistency," in *Advances in Experimental Social Psychology*, Vol. 14, ed. Leonard Berkowitz, New York: Academic Press, 162-202.
- Feldman, Jack and John G. Lynch (1988), "Self-generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (August), 421-435.
- Fischhoff, Baruch (1991), "Value Elicitation: Is There Anything in There?" *American Psychologist*, 46 (August), 833-847.
- , Paul Slovic, and Sarah Lichtenstein (1980), "Knowing What You Want: Measuring Labile Values," in *Cognitive Processes in Choice and Decision Behavior*, Vol. 7, ed. Thomas S. Wallsten, Hillsdale, NJ: Erlbaum, 117-141.
- Fishbein, Martin and Icek Ajzen (1975), *Belief, Attitude, Intention, and Behavior*, Reading, MA.: Addison-Wesley.
- Greenwald, Anthony G., Catherine G. Carnot, Rebecca Beach, and Barbara Young (1987), "Increased Voting Behavior by Asking People if They Expect to Vote," *Journal of Applied Psychology*, 72 (2), 315-318.
- Gregory, William L., Robert B. Cialdini, and Kathleen M. Carpenter (1982), "Self-relevant Scenarios as Mediators of Likelihood Estimates and Compliance: Does Imagining Make It So?" *Journal of Personality and Social Psychology*, 43 (July), 89-99.
- Hardie, Bruce, Eric Johnson, and Peter Fader (1992), "Modeling Loss Aversion and Reference Dependence Effects on Broad Choice," working paper, Marketing Department, Wharton School, University of Pennsylvania, Philadelphia, PA 19104.
- Hippler, Hans-J., Norbert Schwarz, and Seymour Sudman (1987), *Social Information Processing and Survey Methodology*, New York: Springer.
- Hirt, Edward R. and Steven J. Sherman (1985), "The Role of Prior Knowledge in Explaining Hypothetical Events," *Journal of Experimental Social Psychology*, 21 (November), 519-543.
- Hoch, Stephen J. (1984), "Availability and Interference in Predictive Judgment," *Journal of Experimental Psychology*, 10 (4), 649-662.
- Infosino, William (1986), "Forecasting New Product Sales from Likelihood of Purchase Ratings," *Marketing Science*, 5 (Fall), 372-384.
- Kahneman, Daniel and J. Snell (1990), "Predicting Utility," in *Insights in Decision-Making: A Tribute to Hillel J.*

- Einhorn, ed. Robin M. Hogarth, Chicago: University of Chicago Press, 295-310.
- and Amos Tversky (1982), "The Simulation Heuristic," in *Judgment under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman et al., New York: Cambridge University Press, 201-208.
- Kalwani, Manohar U. and Alvin J. Silk (1982), "On the Reliability and Predictive Validity of Purchase Intention Measures," *Marketing Science*, 1 (Summer), 243-286.
- Kardes, Frank R., Chris T. Allen, and Manuel J. Pontes (1993), "Effects of Multiple Measurement Operations on Consumer Judgment: Measurement Reliability or Reactivity?" in *Advances in Consumer Research*, Vol. 20, ed. Leigh McAlister and Michael L. Rothschild, Provo, UT: Association for Consumer Research, in press.
- Koehler, Derek J. (1991), "Exploration, Imagination, and Confidence in Judgment," *Psychological Bulletin*, 110 (3), 499-519.
- Lehmann, Donald R. (1989), *Market Research and Analysis*, Homewood, IL: Irwin.
- Lutz, Richard J. (1991), Editorial, *Journal of Consumer Research*, 17 (March).
- MacKenzie, Scott B., Richard J. Lutz, and George Belch (1986), "The Role of Attitude toward the Ad as a Mediator of Advertising Effectiveness: A Test of Competing Explanations," *Journal of Marketing Research*, 23 (May), 130-143.
- McNeil, John M. (1974), "Federal Programs to Measure Consumer Purchase Expectations, 1946-73: A Post-mortem," *Journal of Consumer Research*, 1 (December), 1-10.
- Morwitz, Vicki G. (1993), "The Predictive Validity of Time Conditioned Intent: Will I Buy and When Will I Buy?" working paper, Marketing Department, New York University, New York, NY 10012.
- Payne, John W., James R. Bettman, and Eric J. Johnson (1992), "Behavioral Decision Research: A Constructive Processing Perspective," *Annual Review of Psychology*, 43, 87-131.
- Schmittlein, David C. and Kristiaan Helsen (1988), "Analyzing Duration Times in Marketing Research," Working Paper no. 88-012, Wharton School, University of Pennsylvania, Philadelphia PA 19104.
- Sewall, Murphy A. (1978), "Market Segmentation Based on Consumer Ratings of Proposed Product Design," *Journal of Marketing Research*, 15 (November), 557-564.
- Sheppard, Blair H., John Hartwick, and Paul R. Warshaw (1988), "The Theory of Reasoned Action: A Meta-analysis of Past Research with Recommendations for Modifications and Future Research," *Journal of Consumer Research*, 15 (December), 325-343.
- Sherman, Steven J. (1980), "On the Self-erasing Nature of Errors of Prediction," *Journal of Personality and Social Psychology*, 39 (August), 211-221.
- , Kim S. Zehner, James Johnson, and Edward R. Hirt (1983), "Social Explanation: The Role of Timing, Set, and Recall on Subjective Likelihood Estimates," *Journal of Personality and Social Psychology*, 44 (June), 1127-1143.
- Silk, Alvin J. and Glen L. Urban (1978), "Pre-test Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *Journal of Marketing Research*, 15 (May), 171-191.
- Sudman, Seymour and Robert Ferber (1979), *Consumer Panels*, Chicago: American Marketing Association.
- Tesser, Abraham (1978), "Self-generated Attitude Change," in *Advances in Experimental Social Psychology*, Vol. 11, ed. Leonard Berkowitz, New York: Academic Press, 289-338.
- Tourangeau, Roger and Kenneth A. Rasinski (1988), "Cognitive Processes Underlying Context Effects in Attitude Measurement," *Psychological Bulletin*, 103 (May), 299-314.