Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 23 (2007) 347–364
international journal of forecasting
www.elsevier.com/locate/ijforecast

When do purchase intentions predict sales?

Vicki G. Morwitz [a,*], Joel H. Steckel [a], Alok Gupta [b]

[a] New York University, United States
[b] Rosetta Marketing Strategies Group, United States

Abstract

Marketing managers routinely use purchase intentions to predict sales. The purpose of this paper is to identify the factors associated with an increased or decreased correlation between purchase intentions and actual purchasing. Using two studies, we examine the data collected from a wide range of different settings which reflect the real world diversity in how intentions studies are conducted. The results indicate that intentions are more correlated with purchases: 1) for existing products than for new ones; 2) for durable goods than for non-durable goods; 3) for short than for long time horizons; 4) when respondents are asked to provide intentions to purchase specific brands or models than when they are asked to provide intentions to buy at the product category level; 5) when purchases are measured in terms of trial rates than when they are measured in terms of total market sales; and 6) when purchase intentions are collected in a comparative mode than when they are collected monadically.

Keywords: Purchase intentions; Sales forecasting; Meta-analysis; Marketing research

1. Introduction

Marketing managers routinely use purchase intentions data to make strategic decisions concerning both new and existing products, and the marketing programs that support them. For new products, purchase intentions are used in concept tests to help managers determine whether a concept merits further development, and in product tests to direct attention to whether a new product merits being launched. Furthermore, in planning the launch of a new product, purchase intentions help the manager decide in which geographic markets and to which customer segments the product should be launched (Sewall, 1978; Silk & Urban, 1978; Urban & Hauser, 1993). For existing products, purchase intentions are used to forecast future demand (Juster, 1966; Morrison, 1979). These forecasts are useful inputs when making decisions, such as whether to increase or reduce production levels, whether to change the size of the sales force, and whether to initiate a price change. In addition, purchase intentions are used to pretest advertising and evaluate proposed promotions for both new and existing products (Bird & Ehrenberg, 1966). Purchase intentions are also extensively used by academic researchers as proxy measures for purchase behavior (e.g. Schlosser, 2003).

* Corresponding author. Stern School of Business, New York University, 40 West 4th Street, Suite 807, New York, NY 10012, United States. Tel.: +1 212 998 0518; fax: +1 212 995 4006.
E-mail address: vmorwitz@stern.nyu.edu (V.G. Morwitz).

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2007.05.015

When managers and academic researchers rely on purchase intentions, they hope, and implicitly assume, that these measures will be predictive of subsequent purchases. This notion is a cornerstone of many theoretical models of consumer behavior. For example, Fishbein and Ajzen (1975, p. 368–369) wrote, "if one wants to know whether or not an individual will perform a given behavior, the simplest and probably the most efficient thing one can do is to ask the individual whether he intends to perform that behavior." According to Bagozzi (1983, p. 145) "intentions constitute a willful state of choice where one makes a self-implicated statement as to a future course of action." Warshaw (1980) notes that most formal consumer behavior models show intent as being an intervening variable between attitude and choice behavior, implying that intentions outperform beliefs or other cognitive measures as behavioral correlates (e.g. Engel, Blackwell, & Kollat, 1978; Howard & Sheth, 1969).

Unfortunately, the signal from empirical investigations of the link between respondents' stated intentions and their ultimate behavior is not as clear. While most studies find a significant positive relationship between intent and behavior (Bemmaor, 1995; Clawson, 1971; Ferber & Piskie, 1965; Granbois & Summers, 1975; Newberry, Kleinz, & Boshoff, 2003; Pickering & Isherwood, 1974; Taylor, Houlahan, & Gabriel, 1975), the strength of this relationship seems to vary quite a bit. For example, in a meta-analysis of a wide range of applications of the Fishbein and Ajzen model, Sheppard, Hartwick, and Warshaw (1988) found that the frequency weighted average correlation for the intention–behavior relationship was 0.53; however, there were substantial variations in the correlations across the studies they examined. Indeed, the 95% confidence limits of the average correlation were 0.15 and 0.92.

A natural question to ask, then, is "why do these correlations vary so much?" The theory of planned behavior states that intentions should only predict behavior if the intentions are measured just prior to the performance of the behavior, and if the behavior is under the individual's sole volitional control (Ajzen, 1985). However, in many marketing research studies these conditions are difficult to meet. A typical study may involve exposing respondents to a new concept description (e.g., a new automobile whose design is being considered by an automobile producer) and measuring both their attitude toward the concept and their intentions to purchase it in the future. The respondents' intentions may change between the time of the survey and the time of a subsequent actual purchase decision. In addition, a respondent will provide his or her own intention to purchase the product, but other individuals in the respondent's household may also play a role in the final purchase decision.

The objective of this research, therefore, is to identify the factors associated with an increased or decreased correlation between purchase intentions and actual purchasing. In two studies, we examine data collected from a wide range of different settings that reflect the real world diversity in how intentions studies are conducted. In the next section we develop some *a priori* hypotheses concerning the factors that moderate the intent–behavior relationship. We then describe and report the results from a meta-analysis conducted in Study 1, and examine the results from Study 2 (a second data set of sixty product tests). Finally, we discuss the implications of our results, the limitations of this work, and opportunities for continued research.

2. Potential moderators of the relationship between purchase intentions and behavior

In Study 1, we used a meta-analysis (Assmus, Farley, & Lehmann, 1984; Hunter, Schmidt, & Jackson, 1982) to examine factors that moderate the correlation between purchase intentions and purchase behavior. Assmus et al. (1984) suggest the use of the following three categories of moderators in meta-analyses: 1) differences in the research environment, 2) differences in measurement, and 3) differences in estimation. They also suggest a fourth category, namely differences in model specification, and this is relevant for the many studies that have used meta-analyses to examine differences in the estimated parameters of a theoretical model. However, this category is not relevant for our research since we are examining a general summary statistic (i.e., a correlation coefficient), and not the estimated parameters of a specific model.

We examine four dimensions related to the specifics of the research environment: i) the type of product (new versus existing, durable versus non-durable), ii) the level of product specificity for which consumers were asked to provide intentions (brand level, sub-brand (variants/flavors) level, or product category level), iii) the type of study (experiments versus surveys, academic versus commercial), and iv) the type of respondents (consumer versus business). We consider three dimensions related to measurement: i) how intentions were measured, ii) how behavior was measured, and iii) the length of time between the two measurements. Finally, we consider one dimension related to estimation: how the relationship between intentions and behavior (i.e., the correlation coefficient) was computed (e.g. across products or across respondents). Whenever possible, we advance hypotheses about specifically how we expect these variables to moderate the intent–behavior relationship.

2.1. Specifics of the research environment

2.1.1. Type of product studied

Although no studies have systematically examined how the correlation between purchase intentions and behavior varies with the type of product, prior studies provide some evidence suggesting that the nature of the intentions–behavior relationship is different for different types of products (Bemmaor, 1995; Bird & Ehrenberg, 1966; Ferber & Piskie, 1965; Granbois & Summers, 1975; Jamieson & Bass, 1989; Kalwani & Silk, 1982). For example, Bemmaor's (1995) results suggest that intentions predict behavior for existing consumer products, but not for new products or products targeted at business markets. Kalwani and Silk (1982) found that a linear model provided a good description of the relationship between intent and purchases for durable goods (i.e., the higher the intention, the more likely the individual is to purchase), but that a piecewise linear model provided a better fit for non-durable or packaged goods. Ferber and Piskie (1965) found that consumers were more likely to fulfill intentions for services than for durable goods.

Based on this literature, we therefore examine how the strength of the intent–behavior relationship varies across two product-related dimensions: (1) whether the product is new or existing, and (2) whether the product is a durable or non-durable good. For each dimension we first define the different product types, and then develop hypotheses. We suggest that the nature of the intention–behavior relationship varies with the type of product because the consumers' levels of familiarity and product knowledge, and their involvement in the decision processes, vary with the type of product under consideration.

2.1.2. New versus existing products

There are several different definitions of what is meant by a "new product" (Booz, Allen, & Hamilton, 1982). A product might be a variant of an existing product (e.g., a new flavor) that is new to the firm, but not to the market. Alternatively, the product might be a variant of an existing product that is new to both the firm and the market. Finally, a product might have a product form that is new to the market. We consider a product to be new if it is new to the market, i.e. it has at least one differentiating physical attribute that no existing competitor has. We view the introduction of such an attribute as triggering a change in the way that customers evaluate the product.

Consumers are likely to have greater familiarity with, and knowledge of, existing products than new products (Johnson & Russo, 1984). For existing products, consumers have the opportunity to obtain product information by examining the product and any available marketing communications such as advertisements, by talking to salespeople and other consumers who have had experience with the product, by reading about the product in newspapers and magazines, and by retrieving their own memories of product related experiences. In contrast, for new products, consumers are typically asked to state their purchase intentions after they have been exposed to a concept statement (a limited description of the product), or after having been exposed to the product during a product test, such as a taste test. Overall, we expect consumers to have more detailed information available for forming intentions to purchase existing products compared to new products. It should therefore be easier for consumers to accurately predict whether or not they will purchase an existing product than a new product.

In addition, prior research in social psychology has shown that the relationship between intentions and behavior is stronger if the behavior is one for which individuals have prior experience (Bentler & Speckart, 1979; Feldman & Lynch, 1988). Consumers are more likely to have prior experience with existing products than new ones. Based on these arguments, we expect that the correlation between purchase intentions and behavior will be higher for existing products than for new products.

H1. The correlation between purchase intentions and behavior will be higher for existing products than for new products.

2.1.3. Durables versus non-durables

According to Kotler and Keller (2006, p. 324), durable goods are "tangible goods that normally survive many uses" and non-durable goods are "tangible goods that are normally consumed in one or a few uses." Two characteristics of consumer durables are likely to affect the predominant decision process. First, durables are used repeatedly; second, they tend to be higher in price than non-durables. The potential (negative) consequences of making an incorrect purchase decision may therefore be large for durable goods. In contrast, non-durable goods are usually lower in price and are typically consumed over much shorter time periods than durables. Mistakes are easily corrected at the next purchase occasion in the product category. All other things being held constant, purchase decisions for durable goods are seen as more important to the consumer than purchase decisions for non-durable goods. Such decisions are often characterized by having high degrees of involvement (Antil, 1984).

It is likely that consumers spend more time gathering information and evaluating alternatives when they consider purchasing high involvement durable goods than when they consider purchasing low involvement non-durable goods. We expect that consumers should be better able to predict their own future behavior for decision making contexts that are more involving than for contexts that are less involving, because of the more formal nature of the decision process (Olshavsky & Granbois, 1979). In addition, the purchase behavior for non-durables can be considered to be more random because there is more variation in consumers' tastes (e.g., variety seeking, changes in tastes over time, etc.), and at least in some cases there will be more variation in the purchase environment for non-durables (more competitors, more frequent use of promotions, etc.)[1] Therefore, we expect the intent–behavior relationship to be stronger for durable goods than for non-durable goods.

H2. The correlation between purchase intentions and behavior will be higher for durable goods than for non-durable goods.

2.1.4. Product level

Purchase intentions are measured at different levels of product specificity (i.e., consumers are asked about their intentions to purchase models, brands, or categories). For example, many commercial studies ask consumers to provide their intentions for various brands or models of a brand (e.g. different flavors of yogurt). Many academic studies, on the other hand, ask respondents about their intentions to buy at the product category level (e.g., Jamieson & Bass, 1989; Pickering & Isherwood, 1974). For the studies included in our database, respondents were sometimes asked whether they intended to buy a flavor or model of a single brand; other times they were asked whether they intended to buy a brand of product within in a single product category; and still other times they were asked whether they intended to purchase any product within a specific product category.

We have no *a priori* expectations about whether or not and how the differences in product level would affect the intent–behavior relationship. Therefore, we do not posit hypotheses, and instead examine any differences in an exploratory fashion.

2.1.5. Type of study

Intentions data are collected for a variety of different types of studies. Businesses collect intentions measures during concept and product tests, and as part of other research projects (e.g., tracking surveys, satisfaction surveys). Academics also collect intentions measures in experimental and survey research. Our database contains studies for five different study types: concept tests, product tests, commercial surveys, academic experiments, academic surveys.

We have no *a priori* expectations about whether or not and how different study types would affect the intent–behavior relationship. Therefore, we do not posit hypotheses, and again examine any differences in an exploratory fashion.

2.1.6. Customer type

Intentions data are used in both consumer and business-to-business marketing. Traditionally, buyer behavior in the business-to-business arena has been thought to be more rational and scientific than in the consumer arena. Business-to-business purchases are more likely to be made with objective criteria relating to economic and technical considerations than consumer purchases are. This is not meant to imply that business customers behave in a manner devoid of emotion and personal prejudice. Rather, the professional aspects and

[1] We thank an anonymous reviewer for this comment.

personal accountability inherent in business-to-business buying impose more of a rational and deliberate approach than that in the typical consumer purchase (Ames & Hlavacek, 1984, Chapter 3).

Given the higher degree of rationality and deliberate processes for business versus consumer respondents, we posit that the stated intentions–behavior correlation will be higher for business customers than for consumers.

H3. The correlation between purchase intentions and behavior will be higher for business customers than for consumers.

2.2. Measurement

The second major category of potential intention–behavior moderators involves the measurement of these constructs. Studies differ with respect to intent measurement, behavior measurement, and the length of the time horizon between the intent and behavior measurements.

2.2.1. Intent measurement

Measures of intent differ in the precise question wording, the number of scale points for the response, and how responses are summarized. Questions can be worded using "likelihood of purchase" or "likelihood of trial" terms. A specific time horizon may or may not be included in the questions (i.e. "How likely are you to buy within the next six months?" e.g. Morwitz & Schmittlein, 1992). Intentions are often measured using binary (Kalish & Soref, 1993; Manski, 1990), five-point (Taylor et al., 1975) or eleven-point scales (Juster, 1966). In addition, intentions data are often summarized and reported in different ways. Some studies report the mean intention score (Miniard, Obermiller, & Page, 1982; Miniard, Obermiller, & Page, 1983; Warshaw, 1980), others report the median (Sewall, 1981), and others report the proportion of respondents who checked off the "top box" of the intention scale (Pringle, Wilson, & Brody, 1982).

In principle, any or all of these differences can affect the strength of the intention–behavior relationship. We examine how the number of scale points and the summary statistic used to report intentions moderate the strength of the intent–behavior relationship. Unfortunately, the precise question wording is missing from descriptions of many of the studies, and relatively few specify whether or not the purchase intention question was asked with a time horizon.

2.2.2. Number of scale points

Kalwani and Silk (1982) show that if stated intentions are distributed as a beta binomial, then the reliability of an intentions scale increases with the number of scale points. However, they also discuss how the presence of response style biases may increase with the number of scale points. The presence of these biases is likely to reduce the strength of the intent–behavior relationship. Overall, Kalwani and Silk conclude that "the relative merits of longer versus shorter intentions scales remains an unresolved question" (p. 264). Therefore, we do not posit any specific hypotheses concerning how the number of scale points moderates the intent–behavior relationship, but rather examine this in an exploratory fashion.

2.2.3. Intention summary

In addition to the number of points in the intentions scale, studies may differ in the summary statistic used to describe intentions (e.g., mean intent, median intent, percent of respondents in the top intent scale point, etc.). For many distributions, these different measures are likely to be positively correlated. Given this, the strength of the relationship between intentions and behavior might not vary much with the intention summary statistic. On the other hand, one might argue that measures that use all of the available scale information (e.g., a mean versus whether or not the top scale point was checked) should be more predictive. We therefore examine this in an exploratory fashion, and do not posit a specific hypothesis.

2.2.4. Behavior measurement

Studies also differ in the summary statistic used to measure purchase behavior. In some studies the behavior measures reflect the trial rates (Jamieson, 1986; Warshaw, 1980; Wells, 1961), while in other studies the behavior measures reflect the aggregate sales or the share data (Harris, 1964; Sewall, 1981; Silk & Urban, 1978). The latter class includes most commercial applications.

To the best of our knowledge, there has been no empirical work that has examined how the intent–behavior relationship varies depending on whether behavior is measured with trial, share, or market sales.

However, we expect intentions to be better predictors of the percentage who will buy a product than market sales. Market sales for a product reflect not only whether or not consumers were willing to try the product, but also consumers' repeat purchasing decisions and purchase quantity decisions. Since we expect that consumers will be better able to predict whether they would try a product once than whether they would buy it more than once and how often they would buy it, we predict that the intent–behavior correlation will be stronger when behavior is measured as the percentage who purchase, than when it is measured as market sales.

H4. The correlation between purchase intentions and behavior will be higher when behavior is measured as the percentage who purchase than when it is measured as market sales.

2.2.5. Length of the interval between intent and behavior measurement

The strength of the relationship between intentions and behavior might vary with the length of time between when intentions and behavior are measured. Fishbein and Ajzen (1975) state that purchase intentions can only predict future behavior when they are measured just prior to the performance of the behavior. This theory suggests that the strength of the relationship between intent and behavior should decrease as the time between measurements increases. Common sense also dictates that people should be better able to predict what they will do in the short-run than in the long-run. One reason for the decreasing ability of intentions to predict behavior through time is that new information about factors that affect the behavior may become available after the elicitation of intentions (Manski, 1990). For example, during the 18 months Campbell Soup spent developing a product called Juiceworks, three competitor products were introduced (Miller & Tsiantor, 1987). In a situation such as this, it is not surprising that the intentions measured during an early concept test may not adequately predict the ultimate market performance.

However, some empirical evidence also suggests that the strength of the relationship between intentions and behavior might increase with time. Research on the planning fallacy has shown that people underestimate the length of time it will take them to complete a task (Buehler, Griffin, & Ross, 1994). Related to this, research has shown in a consumer context that consumers sometimes delay purchasing a product relative to their original purchase plans (Greenleaf & Lehmann, 1995). For example, Morwitz (1995) showed that consumers who are making a first time purchase of a personal computer underestimate how many months it will take them to actually purchase the computer. In situations where consumers are likely to delay making a purchase, the relationship between intentions and behavior might increase with the amount of time between intent and behavior measurement because consumers will have more time to fulfill their purchase intentions.

Since it is not clear whether the strength of the intent–behavior relationship should increase or decrease with the length of time between the intent and behavior measurement, we do not posit any formal hypothesis, and examine the role of time in an exploratory fashion.

2.3. Estimation

Finally, we examine how the strength of the intent–behavior relationship depends on the precise manner in which the correlation between them is computed. This depends on both the amount of information available from each study and the number of products examined in the study.

2.3.1. Correlation form

Managers often compare measures of purchase intentions or other related measures for a focal product to benchmark measures obtained from other products in order to assess the relative demand for the focal product (Shocker & Hall, 1986; Urban & Katz, 1983; Wind, 1982). Managers should only make decisions based on these comparisons if the correlation between intentions and behavior across products is high. Since managers often interpret an intentions measure for one product by comparing it to comparable measures for other products, and since a primary goal of this paper is to help managers design and interpret intentions studies, we computed the intent–behavior correlation across products whenever possible. In particular, when three or more products are evaluated within a single study, we compute an across-product correlation. That is, we take the correlation between an aggregate intent measure and an aggregate behavior measure across products. When fewer than three products are evaluated in a single study we cannot compute a correlation across products, and we instead computed an across-respondent correlation. Studies that report data for fewer than three products were not included in our database unless they provided enough individual level data to allow us to compute a point-biserial correlation across respondents. The point-biserial correlation is between each respondent's stated level of intent and whether or not each respondent eventually purchased the product. Note that this correlation is computed across respondents, and within a product.

Correlations between probabilistic (but essentially continuous) predictions, such as those implied by stated intentions and binary outcomes, such as buy–don't buy, have upper bounds which are considerably less than one (Hunter & Schmidt, 1990; Morrison, 1972). On the other hand, correlations between probabilistic predictions and continuous outcomes, such as penetration rates, are not similarly encumbered. Therefore, we expect that when the intention–behavior correlation is computed across products it will be higher than when it is computed across respondents (i.e., point-biserial correlations).

H5. The correlation between purchase intentions and behavior will be higher when it is computed across products than when it is computed across respondents.

3. Study 1: meta-analysis

3.1. Data included in the meta-analysis

The first step in a meta-analysis involves selecting the studies to be included. We collected studies from two classes of sources: commercial and academic. The commercial data were collected with the aid of a Marketing Science Institute solicitation, along with a few personal contacts, and are proprietary; however, we were given permission to describe them in terms of the factors examined in this paper. The academic data were obtained by performing literature searches in marketing, social psychology, economics and statistics journals over the period 1940–2006. We included any relevant study that measured both purchase intentions and purchase behavior, and either reported an intention–behavior correlation or provided raw data in enough detail for us to compute such a correlation. Note that there were also many intentions–behavior studies in non-marketing disciplines that did not relate specifically to purchasing behavior, and these were not included. In addition, we do not include data from simulated test markets such as ASSESSOR, where consumers actually shop for products in a simulated store, because consumers' responses are closer to actual purchase than stated purchase intentions.

Our efforts produced a database which includes intentions and behavior data from 40 different studies conducted between the years of 1957 and 2006. Combined, these studies provide data from more than 65,000 consumers on more than 200 different products. The relevant details of the studies are presented in Table 1. The average correlation between purchase intentions and behavior across these 40 studies is .49 and the standard deviation is .31. The correlations range from a low value of −.13 to a high value of .99. Although on average the correlation between intentions and behavior was positive, as we would expect, we note that there were a few studies for which there appears to be no relationship, or even a slightly negative relationship, between intentions and behavior. Similar results were found by Sheppard et al. (1988) in their meta-analysis of applications of the theory of reasoned action.

Each study in the meta-analysis was classified on the basis of the factors described in the previous section. Thus, for each study we have a value for the intention–behavior correlation, and we know how that study would be classified for each factor of interest. Table 2 lists the number of studies at each level for each factor, and when we have a prediction, also summarizes the expected effect each factor will have on the strength of the intent–behavior relationship.

3.2. Meta-analysis methodology

A common approach when conducting a meta-analysis is to configure the study categories into a natural quasi-experimental design and perform a dummy variable regression (e.g. Assmus et al., 1984; Farley, Lehmann, & Ryan, 1981; Sultan, Farley, & Lehmann, 1990). This allows the researcher to both assess the impact of each factor by examining the statistical significance of the regression coefficients, and establish a baseline prediction for the outcome of any specific or hypothetical study by applying the regression equation to the dummy variables which categorize the study of interest.

Table 1
Description of studies included in the meta-analysis (Study 1)

Study	R	New/ existing	Durable/ non-durable	Product level	Study type	Customer type	Number of scale points	Intention summary	Behavior measure	Time between intent and behavior	Correlation form
Proprietary	.921	New	Non-durable	Model	Concept test	Consumer	5–7	Top box	Pen/perc	>12 months	Product
Proprietary	.671	New	Non-durable	Model	Concept test	Consumer	5–7	Top box	Sales	12 months	Product
Proprietary	.659	Existing	Durable	Category	Commercial survey	Business	Other	Med/mean	Sales	6 months	Product
Proprietary	.588	Existing	Non-durable	Brand	Product test	Consumer	5–7	Top box	Sales	12 months	Product
Proprietary	.496	New	Non-durable	Model	Product test	Consumer	5–7	Top box	Sales	12 months	Product
Proprietary	.452	Existing	Non-durable	Model	Concept test	Consumer	5–7	Top box	Pen/perc	dk	Product
Proprietary	.436	New	Non-durable	Model	Concept test	Consumer	Other	Top box	Sales	12 months	Product
Proprietary	.301	New	Non-durable	Model	Product test	Consumer	5–7	Top box	Sales	12 months	Product
Proprietary	.233	New	Non-durable	Model	Concept test	Consumer	5–7	Top box	Sales	12 months	Product
Proprietary	.178	New	Non-durable	Model	Product test	Consumer	5–7	Top box	Sales	12 months	Product
Proprietary	.142	New	Non-durable	Model	Concept test	Consumer	5–7	Top box	Sales	12 months	Product
Proprietary	−.093	New	Non-durable	Brand	Product test	Consumer	5–7	Top box	Sales	12 months	Product
Wells (1961)	.992	Existing	Non-durable	Brand	Commercial survey	Consumer	5–7	Top box	Pen/perc	under 1 month	Product
Warshaw (1980)	.975	Existing	Non-durable	Brand	Academic survey	Consumer	11–21	Med/mean	Pen/perc	under 1 month	Product
Wells (1961)	.969	Existing	Non-durable	Brand	Commercial survey	Consumer	5–7	Top box	Pen/perc	under 1 month	Product
Pickering and Isherwood (1974)	.929	Existing	Durable	Category	Academic survey	Consumer	11–21	Top box	Pen/perc	>12 months	Product
Harris (1964)	.910	New	Durable	Model	Academic experiment	Consumer	11–21	Top box	Sales	6 months	Product
Kalish and Soref (1993)	.903	Existing	Durable	Category	Commercial survey	Business	Other	Top box	Pen/perc	12 months	Product
Wilson, Mathews, and Harvey (1975)	.900	Existing	Non-durable	Brand	Academic survey	Consumer	5–7	Level	Other	dk	Individual
Miniard et al. (1983)	.873	Existing	Non-durable	Brand	Academic experiment	Consumer	11–21	Top box	Pen/perc	Under 1 month	Product
Miniard et al. (1982)	.808	Existing	Non-durable	Brand	Academic experiment	Consumer	11–21	Med/mean	Pen/perc	Under 1 month	Product
Theil and Kosobud (1968)	.703	Existing	Durable	Category	Commercial survey	Consumer	5–7	Level	Pen/perc	12 months	Individual
Gormley (1974)	.550	Existing	Non-durable	Brand	Academic survey	Consumer	5–7	Top box	Other	Under 1 month	Product
McQuarrie (1988)	.513	Existing	Durable	Category	Academic survey	Consumer	Other	Level	Other	>12 months	Individual
Juster (1966); Morrison (1979)	.509	Existing	Durable	Category	Commercial survey	Consumer	11–21	Level	Pen/perc	>12 months	Individual
Bemmaor (1995)	.508	Existing	Durable	Category	Commercial survey	Consumer	11–21	Top box	Pen/perc	12 months	Product
De Janosi (1959)	.430	Existing	Durable	Category	Academic survey	Consumer	5–7	Level	Other	12 months	Individual

Table 1 (*continued*)

Study	*R*	New/ existing	Durable/ non-durable	Product level	Study type	Customer type	Number of scale points	Intention summary	Behavior measure	Time between intent and behavior	Correlation form
Bonfield (1974)	.402	Existing	Non-durable	Brand	Academic survey	Consumer	5–7	Level	Other	dk	Individual
Juster (1966); Morrison (1979)	.397	Existing	Durable	Category	Commercial survey	Consumer	11–21	Level	Pen/perc	>12 months	Individual
Harrell and Bennett (1974)	.370	Existing	Non-durable	Brand	Commercial survey	Consumer	5–7	Level	Other	dk	Individual
Juster (1966); Morrison (1979)	.301	Existing	Durable	Category	Commercial survey	Consumer	11–21	Level	Pen/perc	>12 months	Individual
Morwitz and Schmittlein (1992)	.298	Existing	Durable	Category	Commercial survey	Consumer	Other	Level	Pen/perc	>12 months	Individual
McQuarrie (1988)	.263	Existing	Durable	Category	Academic survey	Consumer	5–7	Level	Pen/perc	>12 months	Individual
Kalish and Soref (1993)	.260	New	Durable	Category	Commercial survey	Business	Other	Level	Pen/perc	>12 months	Individual
Sewall (1981)	.251	New	Durable	Models	Concept test	Consumer	5–7	Med/mean	Sales	dk	Product
Mueller (1957)	.190	Existing	Durable	Category	Academic survey	Consumer	Other	Level	Other	6 months	Individual
Jamieson (1986)	−.131	New	Durable	Category	Academic survey	Consumer	5–7	Top box	Pen/ percent	6 months	Product
Jamieson (1986)	−.134	New	Non-durable	Category	Academic survey	Consumer	5–7	Top box	Pen/perc	6 months	Product

In applying this general approach to our database, we encounter three difficulties. First, the dependent variable from each study that we focus on is a correlation. When a population correlation is non-zero, as we expect to be the case in this context, the distribution of its sample value can be quite skewed and problematic when statistical methods that assume normally distributed dependent variables are used. We remedy this by applying Fisher's Z-transform ($Z=1/2\{\ln(1+r)-\ln(1-r)\}$) to the sample correlation. Z is approximately normally distributed, and is consequently amenable to standard statistical analyses (Snedecor & Cochran, 1989).

Second, since the studies were conducted independently and not with a future meta-analysis in mind, some factors exhibit little variation, while others may be highly correlated. Third, our analysis requires 22 dummy variables. A regression that contains 22 variables with 40 data points has limited degrees of freedom (Darlington, 1968). To address these three issues, we conduct our meta-analysis in a three-step procedure similar to that of Farley, Lehmann, and Ryan (1982). In the first step, we remove any factors for which 90% or more of the studies were conducted at the same level. Specifically, we do not include customer type in our analysis since 90% of the studies involved consumer respondents. Thus we cannot test Hypothesis 3.

In the second step, we remove any factors that are perfectly correlated with some combination of others. Two of the independent variables, correlation form and intention summary, are highly correlated. In all studies for which the correlation form is "between-products", the intention summary is equal to "level" (i.e., when a correlation is computed across respondents and within a product, the correlation is between the respondents' stated "level" of intention and whether or not they purchased). Since the intention summary measure captures all the information provided by the correlation form measure, plus a bit more, correlation form is

Table 2
Summary of factors examined in Study 1

Factor examined		Number in-sample	Dummy variable	Hypothesized sign[a]
Research context				
Product type				
	Existing	24	EXIST	H1:+
	New	16	*Reference*	N/A
	Durable	19	DUR	H2:+
	Non-durable	21	*Reference*	N/A
Product level				
	Brand	11	PRODUCT1	?
	Category	17	PRODUCT2	?
	Model/flavor	12	*Reference*	N/A
Study type				
	Academic experiment	4	TYPENV1	?
	Academic survey	12	TYPENV2	?
	Commercial survey	12	TYPENV3	?
	Concept test	7	TYPENV4	?
	Product test	5	*Reference*	N/A
Customer type				
	Business	4	BUSCUST	H3:+
	Consumer	36	*Reference*	N/A
Measurement				
Number of scale points				
	11–21 points	10	SCALE1	?
	5–7	22	SCALE2	?
	Other	8	*Reference*	N/A
Intention summary				
	Level	14	INDEPVR1	?
	Mean/median	4	INDEPVR2	?
	% Top box	22	*Reference*	N/A
Behavior measure				
	Percent who bought	20	DEPVAR1	H4:+
	Market sales	13	DEPVAR2	–
	Other	7	*Reference*	N/A
Time between intent and behavior				
	One month or less	7	TIME1	?
	Six months	5	TIME2	?
	12 months	13	TIME3	?
	More than 12 months	10	TIME4	?
	Unknown	5	*Reference*	N/A
Estimation				
Correlation form				
	Across products	26	UNIT	H5:+
	Within product	14	*Reference*	N/A

[a] When we hypothesize that the dummy variable will be positive (negative) we indicate that with a + (−). Note that for each variable there is a default level where no dummy variable is assigned. We use N/A for each default level. When we do not have a specific hypothesis we use the symbol "?".

eliminated from further analysis. Therefore we cannot test Hypothesis 5. Nevertheless, any findings we obtain with respect to intention summary would also have to be interpreted as potentially reflecting correlation form.

In the third step, we take the dummy variables contained in the remaining factors and find a "best subset" of them in order to produce a final (dummy variable regression) model. Farley et al. (1982) used stepwise regression to select a subset of their original set of variables for further scrutiny in their meta-analysis of consumer decision process models. While stepwise regression provides a useful heuristic, its biases have been well documented (McIntyre, Montgomery, Srinivasan, & Weitz, 1983; Vanhonacker, 1983). In order to minimize the biases associated with stepwise regression, we use an approach suggested by Mallows (1973) and recommended by Draper and Smith (1981, p. 341). Mallows developed a statistic, $C(p)$, based on the prediction error of a regression model with p parameters. He argues that regressions with small prediction errors will have small error sums of squares, and values of $C(p)$ approximately equal to p. We use Mallows' summary statistic to select a model containing a subset of the independent variables. We first identify the regression models with the highest R-squared values for each potential number of independent variables (i.e., the best 1-variable model, 2-variable model, etc.). We then select the regression model that has the value of $C(p)$ that is closest to p as the final model. Note that because we had limited degrees of freedom, we did not examine any interactions between the study factors.

For each study in the database, we first noted, or if necessary computed, an intent–behavior correlation. We next classified the characteristics of each study and coded dummy variables for each level of each independent factor. The names of the dummy variables are listed in Table 2. Recall that an N level factor requires $N-1$ dummy variables, with one level of the factor taken as the reference level to which each dummy variable is compared.

Some studies report sufficient detail to allow us to compute intention–behavior correlations in a variety of ways. For example, Pickering and Isherwood (1974) report the distribution of the level of intent and the subsequent purchase data (i.e., proportion buying) within each level of intent for 18 existing durable good product categories. In this case, we could either compute 18 across-respondent correlations or compute a single across-product correlation between intent and behavior across the 18 products. Where we had a choice, we always computed the intent–behavior correlation in the manner most commonly used in the commercial studies, in order to make our results as useful as possible for marketing managers. These studies most often computed cross-product correlations, and typically used top box scores to summarize intentions data. Therefore, the Pickering and Isherwood data appear in our database as a single across-products correlation of top box scores with the proportion buying. Several of the studies reported in the academic literature report the same data for different time intervals (i.e., different amounts of time between the intent and behavior measurements). In these cases, we always select the longest time period available.

3.3. Results

Applying the procedure outlined in the previous section, we arrived at the five variable model (six parameters including the intercept) reported in Table 3 ($C(p)=6.122$). All independent variables in the regression are significantly different from zero. The p-values reported in the table are one-sided for the variables for which specific hypotheses are reported in Table 2; the other p-values are two-sided.

3.3.1. Specifics of the research environment

The results indicate that product type and product level both significantly moderate the intent–behavior relationship.

3.3.1.1. Product type. As hypothesized, the dummy variables EXIST (product type = existing) and DUR (product type = durable) were significant in the best subset model. The intention–behavior correlation is significantly higher for existing products than for new ones, supporting Hypothesis 1. It is also significantly higher for durables than for non-durables, supporting Hypothesis 2.

These results have important implications for current practice. They suggest that intentions are least predictive of consumer behavior for the very product types where they are most often used: for new, non-durable products (Gruber, 1970; Haley & Case, 1979; Jamieson & Bass, 1989; Pringle et al., 1982). This highlights the need for improved methods for eliciting purchase intentions for these products, and the need to develop alternative methods for forecasting sales for new concepts and products. Several market research suppliers collect intentions data for clients, and also offer proprietary models that use weighted levels of intent as one input to forecast sales of new packaged goods, (e.g., BASES, ENTRO, ESP, and NEWS, see Morwitz, 1991). These providers report impressive success rates for their proprietary methodologies. An

Table3
Best subset regression of factors affecting the intent–behavior correlation (Study 1)

Factors examined	Dummy variable	Coefficient	Standard error	t	p-value
	INTERCEPT	.3361	.1444	2.328	.0260
Research context					
Product type	EXIST	.3733	.1893	1.972	.0284
	DUR	.4537	.2673	1.697	.0494
Product level	PRODUCT2	−.9579	.3091	−3.099	.0039
Study type					
Customer type					
Measurement					
Number of scale points					
Intention summary					
Behavior measure	DEPVAR1	.5098	.1897	2.688	.0056
Time between intent and behavior	TIME1	.5552	.2554	2.174	.0368
Estimation					
Correlation form					
R^2=.5316	$C(p)$=6.122				

The dependent variable is Fisher's Z transformation of the intent–behavior correlation.

interesting issue, then, is how much weight these models place on intentions data versus the other factors used to forecast new product sales.

3.3.1.2. Product level. The negative coefficient for PRODUCT2 (product level = category) suggests that intention–behavior correlations are significantly lower when the correlation is computed across categories than when it is computed across models, flavors, or brands within a category. Although we did not have any *a priori* hypotheses about this, the result seems logical *post hoc*. It is possible that some of the biases associated with intentions measures may be common to all products within a single product category, but may vary across product categories. For example, consumers may systematically overstate their intentions to purchase socially desirable products (e.g., healthy products, recyclable products), and may systematically understate their intentions to purchase socially undesirable products (e.g., cigarettes, high-fat foods, etc.). Therefore, when a correlation is computed across products within the same product category, the biases associated with the product category are common to all products within the category, and therefore do not affect the strength of the intent–behavior correlation. In contrast, when a correlation is computed across product categories, the biases associated with each category may operate in different directions, and may consequently diminish the strength of the intent–behavior correlation.

3.3.2. Measurement

The results indicate that the type of behavior measure and the time between intent and behavior measurement significantly moderate the intent–behavior relationship.

3.3.2.1. Behavior measurement. Supporting Hypothesis 4, intention–behavior correlations are higher when behavior is measured as the proportion of people that buy, than when behavior is measured as the ultimate market sales. Intentions measures typically ask respondents how likely they are to purchase the product once. It is not surprising, then, that these measures are more highly correlated with behavior measures that reflect trial purchasing than with measures that reflect both trial and repeat purchase behavior.

3.3.2.2. Time between intent and behavior measurement. The positive coefficient for TIME1 (time = one month or less) suggests that shorter intervals are significantly associated with higher intent–behavior correlations. This suggests that consumers are able to more accurately predict their behavior for short time horizons than for longer time horizons.

An alternative explanation is that the time between intent and behavior measurement is correlated with the type of product under examination because it reflects the typical inter-purchase time for different types of products. In particular, we might expect the length of this time interval (denoted LENGTH) to be longer for durable goods than for non-durable goods, since the typical inter-purchase time is longer for durables than for non-durables. We conducted a *t*-test and found that LENGTH was significantly longer for durable goods (14.6) than for non-durable goods (8.3) ($p = .015$). If the TIME1 result was an artifact of the durable–non-durable distinction, the longer LENGTH and higher correlations for durables would produce results directionally opposite to that of a positive, significant coefficient for TIME1. Therefore, we believe that our results suggest that consumers are better able to predict what they will do over short periods of time than long periods of time.

4. Study 2: analysis of sixty comparable product tests

4.1. Study motivation

We conducted Study 2 both to help increase the confidence in the results of Study 1, and to extend the findings. In Study 1 we intentionally examined a wide range of studies with varying characteristics, reflective of the actual diversity in how intentions studies are conducted. However, because the studies included in the meta-analysis were quite varied, and because we could not control for all possible differences across studies, it is possible that the observed results could have been caused by factors not included in the meta-analysis model. Thus, we cannot eliminate the possibility that some factors not enumerated in this study could be masking some differences in the stated intent–behavior relationship and enhancing others. Therefore, in Study 2, we examine data from studies conducted in very similar ways, and attempt to replicate one finding from Study 1 and identify some new ones. To the extent that any findings from the meta-analysis are confirmed in these data, we become more confident in the conclusions we have drawn up to this point in the paper.

4.2. Data and methodology

In Study 2 we analyze data from 60 separate intentions studies obtained from a major multinational packaged goods manufacturer in response to the Marketing Science Institute solicitation. All studies involved tests of actual products that were eventually launched into the market. These data are based on interviews of more than 16,000 consumers and the ultimate sales of these products in the marketplace. These data allow us to re-examine one moderating factor from Study 1 and look at some new ones. In contrast to the studies in the previous analysis, these sixty were conducted in very similar ways. All of the products considered were non-durable goods within the same product category. All of the studies were product tests for different products, conducted at different points in time. For all tests, intentions were measured using a five-point likelihood of purchase scale, and behavior was measured as the twelve-month unit sales adjusted for distribution coverage.

However, there were also some important differences across these studies that enable us to further examine moderators of the intent–behavior relationship. Eight of the product tests involved new product variants, while 18 involved established product variants. (Note that the numbers summed across categories do not add up to 60 because of difficulties in determining whether some of the products tested were new or existing. For similar reasons, the numbers in the groupings below do not always add up to 60.) By comparing the data for these product variants, we can test the validity of our previous finding that intentions are better predictors of sales of existing products than of new products. For twenty of the product tests, respondents only evaluated a single product variant of interest. For 39 of the product tests, respondents were asked to compare the product variant of interest to a different product variant. Finally, in fifty of the product tests, the brand name was not revealed to the respondents prior to intent measurement. In eight, it was. Thus, in addition to further examining the new-existing distinction, we have the ability to determine whether the presence of a brand name and the introduction of comparative testing moderates the intent–behavior relationship.

The exact data provided to us were the intentions value for each respondent and the market sales for each product. For each product we compute the percentage of respondents who checked off the top box of the intentions scale (i.e., “definitely will buy”). This is the aggregate level intentions summary statistic that we examine together with market sales. (We also repeated the analysis using the average intention score for each product, and obtained nearly identical results.) We begin by computing a single intent–behavior correlation across all sixty studies ($r=.414$). The differences between the studies essentially form part of a $2\times2\times2$ factorial with the following treatments: product type (new-existing), brand name (given-not), and mode (comparative-monadic). Unfortunately, the design is extremely unbalanced and some of the cells are empty. In particular, all tests involving new product models were conducted in a comparative mode (and in only one of them was the brand name provided). Thus there are no new product tests done in a monadic mode. Consequently, we must be extremely cautious in interpreting our results. In addition, given the extremely unbalanced nature of the design, rather than using regression analysis as in Study 1, we simply assess the effect of each treatment by computing separate intention–behavior correlations for studies with the same level of each treatment (e.g., we compute one correlation for all studies involving new product variants, and a separate correlation for studies involving existing product variants).

Table 4
Correlations from sixty comparable product tests (Study 2)

Overall	$r=.414$** ($n=60$)
New versus established	
New	$r=-.177$ ($n=8$)
Established	$r=.751$** ($n=18$)
Unknown	$r=.082$ ($n=34$)
Evaluated in comparative or non-comparative mode	
Comparative mode	$r=.530$** ($n=40$)
Non-comparative mode	$r=-.088$ ($n=20$)
Brand name provided or not provided	
Brand name provided	$r=.384$ ($n=8$)
Brand name not provided	$r=.473$** ($n=50$)
Unknown	– ($n=2$)
Comparative mode — new versus established	
New	$r=-.177$ ($n=8$)
Established	$r=.767$** ($n=17$)
Unknown	$r=.068$ ($n=15$)
Brand name not provided — new versus established	
New	$r=-.096$ ($n=7$)
Established	$r=.751$** ($n=18$)
Unknown	$r=-.253$* ($n=25$)

* Significant at $p<.05$.
** Significant at $p<.01$.

4.3. Results

The results are summarized in Table 4. We find further evidence that intentions are better predictors of behavior for existing products than for new products. Specifically, the intent–sales correlation is .751 for the 18 existing products and −.177 for the eight new products. The two correlations are significantly different, and the correlation for existing products is significantly greater than zero, while the correlation for new products does not differ significantly from zero.

We also find that intentions are only significant and positive predictors of behavior for the forty products that were evaluated in a comparative mode ($r=.530$), not for the twenty products that were evaluated in a non-comparative mode ($r=.088$). This finding, which we were unable to pursue in our meta-analysis, is even more salient when we realize that all of the new products were evaluated comparatively. Thus, the comparative correlation might be understated relative to the non-comparative one because new products have inherently lower intention–behavior correlations.

Finally, whether or not respondents are told the brand name of a product does not appear to moderate the intent–behavior relationship. The intent–sales correlation is .384 for the eight products for which the brand name was provided and .473 for the fifty products for which the brand name was not provided. These correlations are not significantly different. Although the correlation was only significantly different from zero when the brand name was not provided, the correlations do not vary as much as for the previous two factors. In addition, the sample size for branded products is quite small ($n=8$). The lack of significance of providing the brand name does not seem to be an artifact of the unbalanced design, since the proportions of new products in each condition are very close (1 out of 8 when the brand name is provided, and 7 out of 50 when the brand name is not provided).

To sum up, this additional data set allowed us to reconfirm the new-existing finding from the meta-analysis, and additionally establish that the intent–behavior correlation is higher when intentions data are collected in a comparative mode than in a non-comparative mode.

5. Conclusion

The analyses described in this paper demonstrate that the strength of the relationship between respondents' stated intentions and their ultimate purchase behavior varies with the types of products that are studied and the way that these data are collected. Specifically, the results suggest that intentions are more correlated with purchases (1) for existing products than for new products; (2) for durable products than for non-durable products; (3) when respondents are asked to provide intentions to purchase specific brands or models than when they are asked to provide intentions to buy at the product category level; (4) when purchase levels are measured in terms of trial rates rather than total market sales; (5) for short time horizons than for long time horizons; and (6) when intentions are collected in a comparative mode than when they are collected monadically.

5.1. Implications

Overall, the results indicate that purchase intentions are predictive of future behavior, and that much of the variation in the intent–behavior relationship can be explained by the characteristics of the study. These results suggest that consumers will be better able to accurately predict their future purchasing when the purchase decision is relatively easy (e.g., the purchase will occur in a short time horizon, the consumer is familiar with and knowledgeable about the product, the product description (level) is explicit, and the tradeoffs involved in purchasing this product versus another are made explicit). They also suggest that intentions will be more predictive of behavior when the consequences of purchasing are great, and consumers therefore deliberate considerably about the purchase decision (e.g., purchasing a high involvement durable good). Future studies should explicitly examine the role of ease of making a purchase decision, the extent to which deliberation is involved in making a purchase decision, and their interactions on the strength of the intent–behavior relationship.

Managers can use the results obtained from the meta-analysis to improve the design of their intentions studies. For example, the results suggest that if intentions are collected in a comparative mode rather than a monadic mode, the correlation between intentions and subsequent behavior will be higher. Similarly, the results suggest that the correlation between intentions and behavior will be greater for short time periods than for longer time periods. However, in some cases managers must include factors in their studies that we have shown reduce the strength of the intent–behavior relationship, because the factors are inherent to what is being studied. In these cases, managers can use these results to help determine how much weight to place on purchase intentions measures in sales forecasts relative to other factors that may be predictive of future sales (e.g., past sales managerial intuition, dollar expenditures on advertising, distribution, promotion, etc.; see Armstrong, Morwitz, & Kumar, 2000, and Kumar, Nagpal, & Venkatesan, 2002, for discussions concerning combining intentions with other predictors of purchasing). For example, although we found that the intent–behavior correlation is higher for existing products than for new products, managers will still want to measure consumers' purchase intentions for new concepts and products to enable them to make strategic decisions about whether and how to launch these products. In these cases, the results of the meta-analysis model presented in Table 3 can be used by managers to determine how strong they should expect the intent–behavior correlation to be, and therefore how much weight they should place on intentions in a forecasting model, given the characteristics of the study. Specifically, an estimate of the intent–behavior correlation can be obtained by multiplying the coefficients in Table 3 by the relevant dummy variables, adding the intercept, and then applying the inverse Fisher's Z transformation shown below (Snedecor & Cochran 1989, p. 475).

For example, a manager who conducted a study where intentions were measured for a new flavor of potato chips (i.e., a new non-durable where intentions data were collected at the product flavor level), should expect a correlation of .324 between intent and dollar sales for the first 12 months ($Z=.336=.3361+.3373(0)+.4537(0)-.9579(0)+.5098(0)+.5552(0)$; $r=.324$). The manager should keep this relatively weak correlation in mind when making strategic decisions based on consumers' stated purchase intentions. Given these results, the manager might decide to place more weight on his or her intuition and on planned expenditures for advertising, distribution, and promotion, and less weight on intentions in a forecasting model. Note that in this case, if the manager could make strategic decisions about this potato chip flavor based on the expected relationship between intentions and the percentage of consumers who would purchase in the first 12 months (versus dollar sales), the expected correlation between intent and behavior would be higher ($Z=.846=.3361+.3373(0)+.4537(0)-.9579(0)+.5098(1)+.5552(0)$; $r=.690$).

5.2. Limitations

Several important limitations should be noted about our research. First, although we did an exhaustive search of the existing literature and several corporations were kind enough to let us look at their records, our overall sample size is still relatively small, forty studies. Purchase intentions have been studied for many years, yet there are still not many documented studies concerning their predictive accuracy.

Second, the studies included in the meta-analysis vary in other dimensions which we were unable to control for in our analyses. While we would have liked to control for all differences across studies, we did not have sufficient information from many of them to be able to control for all of these factors. Overall, though, we believe that the effects of other potential differences across the studies will be smaller in magnitude and less relevant to the marketing manager than the dimensions for which we do control.

Our results are also subject to a truncation bias. The only products for which measures of behavior are available are those that were eventually introduced. Presumably those products are the ones that had higher intention scores. It stands to reason then that a truncation bias would minimize the variance of the intention scores. This would tend to mask the intention–behavior relationship and deflate the correlations that we studied.

For most of the studies examined, intentions and behavior were measured among the same sample of research participants. Chandon, Morwitz, and Reinartz (2005) showed that the relationship between someone's true intentions (their latent intention to buy the product which is independent of the survey measurement) and purchasing will be stronger when both measures are obtained from the same sample. In contrast to the truncation bias, this would tend to strengthen the intention–behavior relationship and enhance the correlations that we studied.

5.3. Future research

The finding that the predictive accuracy of purchase intentions is lower for new products, for non-durable goods, and for more temporally distant purchase occasions suggests the need for future research on methods for improving consumer predictions in these contexts. For low involvement purchases of non-durable goods, respondents may form a purchase intention only once they have been asked in a survey (Morwitz, Johnson, & Schmittlein, 1993). They may therefore not consider all of the aspects of the purchase situation that they would have, had they given the situation more involved thought. They may place less weight on the negative aspects of purchasing the product and place more weight on the positive aspects. For example, they might place more weight on the benefits they expect to obtain by using the product and less weight on the cost of giving up the product they currently use. Wright and Kriewall (1980) demonstrated that consumers who thought that consumption of the product was a distant event used simpler evaluation strategies and placed less weight on negative information. When consumption seemed imminent, more complex evaluation strategies were used and negative information was weighted more heavily. Hoch (1984) found that likelihood judgements for future events were influenced by whichever side of the issue the subjects thought of first. If pro-purchase reasons are thought of first, it might inhibit the generation of anti-purchase reasons and create an overstatement of intent. One way to improve the validity of intention measures might be to prompt the respondent to think of reasons both against and for a future purchase, and to encourage them to imagine the purchase decision as imminent.

Warshaw (1980) suggested that the weak relationship observed between purchase intentions and subsequent purchasing in some previous studies may be the result of the lack of contextually specific intention measures. His research demonstrated that when respondents were asked questions about purchase contexts (e.g., where they would purchase the product), purchase intentions were more predictive of subsequent purchase. The implication is that marketing practitioners should use conditional intentions measures rather than direct measures. However, in a replication of Warshaw's study, Miniard et al. (1983) found no evidence that conditional intention measures are better predictors than direct measures. Given the weaker performance of direct intention measures for new products and non-durable goods observed in this paper, we believe that further studies of the performance of conditional measures should be pursued.

Acknowledgments

The authors thank the Marketing Science Institute for their support of this project, and Bruce Buchanan, Peter Fader and Don Lehmann for their comments on previous drafts.

References

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 11−39). Berlin: Springer–Verlag.

Ames, B. C., & Hlavacek, J. D. (1984). *Managerial marketing for industrial firms.* New York: Random House.

Antil, J. H. (1984). Conceptualization and operationalization of involvement. In T. C. Kinnear (Ed.), *Advances in consumer research*, *Vol. 11* (pp. 203−209). Provo, UT: Association for Consumer Research.

Armstrong, J. S., Morwitz, V. G., & Kumar, V. (2000). Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy? *International Journal of Forecasting*, *16*, 383−397.

Assmus, G., Farley, J. U., & Lehmann, D. R. (1984). How advertising affects sales: Meta-analysis of econometric results. *Journal of Marketing Research*, *21*, 65−74.

Bagozzi, R. P. (1983). A holistic methodology for modeling consumer response to innovation. *Operations Research*, *31*, 128−176.

Bemmaor, A. C. (1995). Predicting behavior from intention-to-buy measures: The parametric case. *Journal of Marketing Research*, *32*, 176−191.

Bentler, P. M., & Speckart, G. (1979). Models of attitude–behavior relations. *Psychological Review*, *86*, 452−464.

Bird, M., & Ehrenberg, A. S. C. (1966). Intentions-to-buy and claimed brand usage. *Operations Research Quarterly*, *17*, 27−46.

Bonfield, E. H. (1974). Attitude, social influence, personal norm, and intention interactions as related to brand purchase behavior. *Journal of Marketing Research*, *11*, 379−389.

Booz, Allen, & Hamilton (1982). *New products management for the 1980s.* New York: Booz, Allen & Hamilton.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the 'planning fallacy': Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*, 366−381.

Chandon, P., Morwitz, V. G., & Reinartz, W. J. (2005). Do intentions really predict behavior? Self-generated validity effects in survey research. *Journal of Marketing*, *69*, 1−14.

Clawson, C. J. (1971). How useful are 90 day purchase probabilities? *Journal of Marketing*, *35*, 43−47.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, *69*, 161−182.

De Janosi, P. (1959). Factors influencing the demand for new automobiles. *Journal of Marketing*, *23*, 412−418.

Draper, N., & Smith, H. (1981). *Applied regression analysis*, 2nd edition. New York: Wiley.

Engel, J., Blackwell, R., & Kollat, D. (1978). *Consumer behavior,* 3rd edition. Hinsdale, Illinois: Dryden Press.

Farley, J. U., Lehmann, D. R., & Ryan, M. J. (1981). Generalizing from imperfect replication. *Journal of Business*, *54*, 597−610.

Farley, J. U., Lehmann, D. R., & Ryan, M. J. (1982). Patterns in parameters of buyer behavior models: Generalizing from sparse replication. *Marketing Science*, *2*, 181−204.

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73*, 421−435.

Ferber, R., & Piskie, R. A. (1965). Subjective probabilities and buying intentions. *Review of Economics and Statistics*, *47*, 322−325.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior.* Reading, MA: Addison–Wesley.

Gormley, R. (1974). A note on seven brand rating scales and subsequent purchases. *Journal of the Marketing Research Society*, *16*, 242−244.

Granbois, D., & Summers, J. O. (1975). Primary and secondary validity of consumer purchase probabilities. *Journal of Consumer Research*, *1*, 31−38.

Greenleaf, E. A., & Lehmann, D. L. (1995). Reasons for substantial delay in consumer decision making. *Journal of Consumer Research*, *22*, 186−199.

Gruber, A. (1970). Purchase intent and purchase probability. *Journal of Advertising Research*, *10*, 23−27.

Haley, R. I., & Case, P. B. (1979). Testing thirteen attitude scales for agreement and brand discrimination. *Journal of Marketing*, *43*, 20−32.

Harrell, G. D., & Bennett, P. D. (1974). An evaluation of the expectancy value model of attitude measurement for physician prescribing behavior. *Journal of Marketing Research*, *11*, 269−278.

Harris, D. (1964). Predicting consumer reactions to product designs. *Journal of Advertising Research*, *4*, 34−37.

Hoch, S. J. (1984). Availability and interference in predictive judgement. *Journal of Experimental Psychology*, *10*, 649−662.

Howard, J. A., & Sheth, J. N. (1969). *The theory of buyer behavior.* New York: John Wiley.

Hunter, J. E., & Schmidt, F. L. (1990). Dichotomization of continuous variables: The implications for meta-analysis. *Journal of Applied Psychology*, *73*, 334−349.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Jamieson, L. F. (1986). *An Investigation of the Relationship Between Purchase Intention and Trial Behavior*. Unpublished Doctoral Dissertation.

Jamieson, L. F., & Bass, F. M. (1989). Adjusting stated purchase intention measures to predict trial purchase of new products: A comparison of models and methods. *Journal of Marketing Research*, *26*, 336−345.

Johnson, E. J., & Russo, J. E. (1984). Product familiarity and learning new information. *Journal of Consumer Research*, *11*, 542−550.

Juster, F. T. (1966). Consumer buying intentions and purchase probability: An experiment in survey design. *Journal of the American Statistical Association*, *61*, 658−696.

Kalish, S., & Soref, E. (1993). On the predictive validity of the intention to purchase measures in business markets. *Working paper 9-1993, Institute for the Study of Business Markets*: Smeal College of Business Administration, Penn State University.

Kalwani, M. U., & Silk, A. J. (1982). On the reliability and predictive validity of purchase intention measures. *Marketing Science*, *1*, 243−286.

Kotler, P., & Keller, K. (2006). *Marketing management: Analysis, planning, and control,* 12th edition. Upper Saddle River: Prentice Hall.

Kumar, V., Nagpal, A., & Venkatesan, R. (2002). Forecasting category sales and market share for wireless telephone subscribers: A combined approach. *International Journal of Forecasting*, *18*, 583−603.

Mallows, C. L. (1973). Some comments on C_p. *Technometrics*, *15*, 661−676.

Manski, C. F. (1990). The use of intentions data to predict behavior: A best-case analysis. *Journal of the American Statistical Association*, *85*, 934−940.

McIntyre, S. H., Montgomery, D. B., Srinivasan, V., & Weitz, B. A. (1983). Evaluating the statistical significance of models developed by stepwise regression. *Journal of Marketing Research*, *20*, 1−11.

McQuarrie, E. F. (1988). An alternative to purchase intentions: The role of prior behavior in consumer expenditure on computers. *Journal of the Market Research Society*, *30*, 407−437.

Miller, A., & Tsiantor, D. (1987, December 28). A test for market research. *Newsweek*, *110*, 32−33.

Miniard, P. W., Obermiller, C., & Page, T. J. (1982). Predicting behavior with intentions: A comparison of conditional versus direct measures. In A. A. Mitchell (Ed.), *Advances in consumer research*, *Vol. 9* (pp. 461−464). Ann Arbor, MI: Association for Consumer Research.

Miniard, P. W., Obermiller, C., & Page, T. J. (1983). A further assessment of measurement influences on the intention–behavior relationship. *Journal of Marketing Research*, *20*, 206−212.

Morrison, D. G. (1972). Upper bounds for correlations between binary outcomes and probabilistic predictions. *Journal of the American Statistical Association*, *67*, 68−70.

Morrison, D. G. (1979). Purchase intentions and purchase behavior. *Journal of Marketing*, *43*, 65−74.

Morwitz, V. G. (1991). *The Link Between Purchase Intentions and Purchase Behavior: Predicting Across Individuals and Over Time*. Doctoral Dissertation, University of Pennsylvania.

Morwitz, V. G. (1995). *Predicting whether and when consumers will purchase from intentions data.* Working paper: Stern School of Business, New York University.

Morwitz, V. G., Johnson, E. J., & Schmittlein, D. C. (1993). Does measuring intent change behavior? *Journal of Consumer Research*, *20*, 46−61.

Morwitz, V. G., & Schmittlein, D. C. (1992). Using segmentation to improve sales forecasts based on purchase intent: Which "intenders" actually buy? *Journal of Marketing Research*, *29*, 391−405.

Mueller, E. (1957). Effects of consumer attitudes on purchases. *American Economic Review*, *47*, 946−965.

Newberry, C. R., Kleinz, B. R., & Boshoff, C. (2003). Managerial implications of predicting purchase behavior from purchase intentions: A retail patronage case study. *Journal of Services Marketing*, *17*, 609−618.

Olshavsky, R. W., & Granbois, D. H. (1979). Consumer decision making: Fact or fiction. *Journal of Consumer Research*, *6*, 93−100.

Pickering, J. F., & Isherwood, B. C. (1974). Purchasing probabilities and consumer buying behavior. *Journal of the Market Research Society*, *16*, 203−226.

Pringle, L. G., Wilson, R. D., & Brody, E. I. (1982). NEWS: A decision-oriented model for new product analysis and forecasting. *Marketing Science*, *1*, 1−28.

Schlosser, A. E. (2003). Experiencing products in the virtual world: The role of goal and imagery in influencing attitudes versus purchase intentions. *Journal of Consumer Research*, *30*, 184−198.

Sewall, M. A. (1978). Market segmentation based on consumer ratings of proposed product design. *Journal of Marketing Research*, *15*, 557−564.

Sewall, M. A. (1981). Relative information contributions of consumer purchase intentions and management judgment as explanators of sales. *Journal of Marketing Research*, *18*, 249−253.

Sheppard, B. H., Hartwick, J., & Warshaw, P. R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research*, *15*, 325−343.

Shocker, A. D., & Hall, W. G. (1986). Pretest market models: A critical evaluation. *The Journal of Product Innovation Management*, *3*, 86−107.

Silk, A. J., & Urban, G. L. (1978). Pre-test market evaluation of new product goods: A model and measurement methodology. *Journal of Marketing Research*, *15*, 171−191.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods,* 8th edition. Ames, IA: Iowa State University Press.

Sultan, F., Farley, J. U., & Lehmann, D. R. (1990). A meta-analysis of applications of diffusion models. *Journal of Marketing Research*, *27*, 70−77.

Taylor, J. W., Houlahan, J. J., & Gabriel, A. C. (1975). The purchase intention question in new product development. *Journal of Marketing*, *39*, 90−92.

Theil, H., & Kosobud, R. F. (1968). How informative are consumer buying intentions surveys? *Review of Economics and Statistics*, *50*, 207−232.

Urban, G. L., & Hauser, J. R. (1993). *Design and marketing of new products,* 2nd edition. Englewood Cliffs, NJ: Prentice–Hall.

Urban, G. L., & Katz, G. M. (1983). Pretest market models: Validation and managerial implications. *Journal of Marketing Research*, *20*, 221−234.

Vanhonacker, W. R. (1983). *Torturing the data with stepwise regression: What will it confess?* Working paper: Graduate School of Business, Columbia University.

Warshaw, P. R. (1980). Predicting purchase and other behaviors from general and contextually specific intentions. *Journal of Marketing Research*, *17*, 26−33.

Wells, W. D. (1961). Measuring readiness to buy. *Harvard Business Review*, *39*, 81−87.

Wilson, D. R., Mathews, H. L., & Harvey, J. W. (1975). An empirical test of the Fishbein intention model. *Journal of Consumer Research*, *1*, 39−48.

Wind, Y. R. (1982). *Product policy: Concepts, methods, and strategy* (pp. 436−461). Reading, MA: Addison–Wesley.

Wright, P., & Kriewall, M. A. (1980). State-of-mind effects on the accuracy with which utility functions predict market place choice. *Journal of Marketing Research*, *17*, 277−293.

Vicki G. Morwitz is a Professor and the Robert Stansky Faculty Research Fellow at New York University.

Joel H. Steckel is Professor at the Marketing Department, Stern School of Business, New York University, 40 W 4th Street, New York, NY 10012.

Alok Gupta is Director at Rosetta Marketing Strategies Group.