A Comparison of Order Fulfillment Methods for Online Ordering

Kenneth K. Boyer
Associate Professor of Operations Management
Eli Broad Graduate School of Management
Michigan State University
East Lansing, MI 48824-1122
Phone: (517) 353-6381
Fax: (517) 432-1112
Email: boyerk@msu.edu

G. Tomas M. Hult
Director, International Business Center (MSU-CIBER) and
Associate Professor of Marketing and Supply Chain Management
Eli Broad Graduate School of Management
Michigan State University
East Lansing, MI 48824-1122
Phone: (517) 353-4336
Fax: (517) 432-1009
Email: hult@msu.edu

December 16, 2003

Keywords: Order Fulfillment; E-commerce; Grocery Home Delivery; Service Quality; Internet Ordering

---

This research was conducted with Grant SES 0216839 from the National Science Foundation. We appreciate the cooperation and assistance of one grocer in the U.S. and two grocers in the U.K.

A Comparison of Order Fulfillment Methods for Online Ordering

ABSTRACT

This study presents an analysis of the growing market for groceries and other foodstuffs ordered via the internet or telephone for delivery to the customer's home. This industry has been growing for the past 5 years at greater than 25% per year while the overall market for foodstuffs has been largely stagnant. The research utilizes data from surveys of over 2,000 customers of five different home delivery grocers. The analysis utilizes two moderating variables (customer experience level and order picking method) and five primary constructs (service quality, product quality, internet quality, time-savings and behavioral intentions). The results indicate that customer perceptions of the primary constructs generally improve as they gain experience with this new method of ordering and receiving groceries. Furthermore, the operational choice of picking method is also shown to have a large impact on customer perceptions – in particular, more experienced customers generally rate the primary constructs higher for DC-based picking than for store-based picking. The study provides support for the hypothesis that direct to customer foodstuffs can be of better freshness, quality and range when picked from a DC because of the ability to shorten the supply chain than from a store. The data suggest that a DC-based picking strategy is viable if grocers can re-shape customer perceptions and master the numerous intricacies of the supply chain.

This research was conducted with a grant from the National Science Foundation. We appreciate the cooperation and assistance of three grocers in the U.S., one grocer in the U.K. and one grocer in Canada.

INTRODUCTION

Groceries are perhaps the most universal commodity, thus competition often spurs supermarkets to go to great lengths to develop new technologies and methods of streamlining both their supply chain and their marketing efforts. Supermarkets are well known as a difficult business to compete in with net profit margins typically about 1-2% of sales. The supply chain challenges associated with supermarkets are enormous: the average supermarket carries 30,000 plus SKUs which are in a constant state of flux and prices that must match the competitor down the street. In an effort to address some of the supply chain challenges, starting in the early 1990s, the grocery industry pursued a major initiative labeled efficient consumer response (ECR) to streamline the supply chain, yet the general consensus has been that changes have been slow and far from successful (Frankel, Goldsby and Whipple, 2002). A 1996 Andersen Consulting study found that manufacturers of packaged goods products spent 13% of sales ($25 billion) on trade promotions and that hidden costs due to supply chain volatility and uncertainty accounted for $5 – 8 billion (Andersen Consulting, 1997). If trade promotions are taken out of the mix, supermarkets across the board lose substantial amounts of money.

During the dot-com mania of the late 1990s and 2000, several start-ups promised to revolutionize the way groceries are bought and sold. Companies like Webvan, Streamline and Homegrocer promised cheaper, more convenient methods of shopping. Yet, Webvan became the standard bearer for dot-bombs by burning through $1 billion in investor capital in a little over a year before going bankrupt in 2001 (Rizzo, 2001). There were numerous problems with the first generation of online grocers, including the challenges of extending the supply chain from existing stores to customer homes and changing customer behavior to embrace a new form of shopping. Clearly these barriers were not met and overcome by Webvan, Streamline and Homegrocer. Yet, despite the relative dearth of publicity, there is evidence that online grocery is alive and growing.

Forrester Research Inc estimates that food and beverage sales on the web will total $3.7 billion for 2003, up 40% from the prior year (Wingfield, 2003). In particular, at least 10 grocers currently offer online ordering for home delivery and all have more than $50 million in sales. These grocers include traditional bricks and mortar stores: Tesco (the world leader with over $600 million in sales in 2002), Sainsbury's, Safeway, Albertsons, PublixDirect and stores that do not have physical stores but choose to deliver from a central distribution center: FreshDirect, Ocado, Grocery Gateway, SimonDelivers and Peapod (Hamilton, 2003; McLaughlin, 2003; Wingfield, 2003). In short, despite reports of its demise there is considerable life in this sector.

The focus of this research is on comparing the two primary operational approaches for picking and delivering customer orders. The most successful online grocers to date, including Tesco, Safeway and Albertsons have all chosen to pick grocery orders at existing stores. This approach has the advantage of minimize the cost of fixed investments, leveraging existing facilities and being close to the customer's home. In contrast, several grocers are bypassing the costs of physical stores and delivering to customers straight from a central distribution center. This approach theoretically can reduce some costs by cutting a link out of the supply chain and also offer fresher produce, meats and dairy items due to the reduced length of the supply chain. The DC-based approach requires large initial investments to build the DC and thus requires large volumes of business. Thus, one of the many reasons for Webvan's failure was an inability to attract enough orders to run the DC at a profit. Current grocers such as FreshDirect, Ocado and Grocery Gateway have learned many lessons from earlier failures and are approaching the problem in new ways. Several of these grocers are approaching profitability and have shown the ability to offer fresher produce through this shortened supply chain (Laseter, Berg and Turner, 2003).

At present, online grocers that pick from existing stores are generally more successful than DC-based grocers, yet there are reasons to believe that cutting existing stores out of the supply

chain offers a solution to many of the grocery industry problems cited above. Since the dot-com bubble burst, most grocers are extremely reticent to release detailed financial or operating details (Mnyandu, 2003). To examine differences in customer perceptions of key aspects of online grocery shopping, we report results from an intensive study that includes a substantial portion of the major players in this growth industry (surveys of over 4000 customers of 7 different companies). In particular, we focus on five grocers in our study that can clearly be classified as employing either a pure store-based picking approach or a pure DC-based picking approach. The two grocers in our study that employ a hybrid approach are excluded from this analysis. This approach allows us to rigorously control for operational differences and evaluate customer perceptions and behaviors.

LITERATURE REVIEW AND HYPOTHESES

Our general research model is shown in Figure 1. The first step in our analysis is to examine two moderating factors: customer experience and picking method. The basic premise of this model holds that experience with online grocery ordering is a critical factor due to the fundamental differences from traditional, in-store shopping. Similarly, the operational choice made by the company to pick orders from either a store or a distribution center will affect numerous customer perceptions. Thus, as shown in Figure 1, we will first examine differences in customer perceptions based on these two moderating factors. The second step in our analysis is to examine relationships between direct factors (service quality, product quality, internet quality, time savings) and behavioral intentions of customers. This will be done using linear regression as moderated by customer experience and picking method. The following sections examine the existing literature for each of the direct factors and develop the specific hypotheses to be tested.

Insert Figure 1 about here

Service Quality

Understanding the impact of service-encounter constructs such as physical good quality, service quality, and the servicescape on behavioral intentions has preoccupied services researchers for more than two decades (Lovelock 1983; Shostack 1977). In addition to the vast amount of support found in service quality literature for this link (e.g., Boulding, Kalra and Staelin, 1999; Cronin and Taylor 1992; Zeithaml, Berry, and Parasuraman 1996), the idea that customers prefer greater service quality is intuitive, particularly if price and other cost elements are held constant. Additionally, equity theory suggests that customers who perceive an organization's delivery of service quality in conjunction with, for example, groceries to be superior are likely to attribute greater equity to the relationship with that organization (Kelley and Davis 1994).

Customer experience level and picking method are both expected to affect perceptions of service quality. First, it appears to be intuitive that customers that experience better service quality will be more loyal and likely to continue purchasing with a given company. However, this relationship is likely to be moderated in an e-commerce setting in which the buying process is dis-intermediated or de-personalized, the importance of the service quality can be expected to change (Kaynama and Black, 2000; Meuter, Ostrom, Roundtree and Bitner, 2000). Second, the picking method should affect service quality by virtue of differences in the operational execution by picking in-store or in a distribution center. In a typology of service organizations, Bitner (1992) highlighted the importance of complexity of the provider's operations as a major factor in customer perceptions of service quality. In the grocery industry in particular, the connection between employees and customers is likely to be less direct for grocers delivering from a distribution center due to the greater average distance between DC and customer than for in-store picking (Yrjola, 2001).

Based on the above arguments, we examine the following hypotheses regarding service quality, customer experience level, picking method and behavioral intentions:

H1(A): Customer perceptions of service quality will differ by customer experience level.

H1(B): Customer perceptions of service quality will differ based on the picking method for selecting items (i.e. store or DC-based).

H1(C): There is an interaction effect between customer experience level and picking method for customer perceptions of service quality.

H2: Service quality is correlated with increased customer satisfaction as measured by behavioral intentions.

Product Quality

Product quality is always an important aspect of the purchasing decision, but the importance generally is intensified when purchasing over the Internet. Numerous researchers have argued that online markets facilitate increased competition and create relatively "friction free" markets (Malone, Yates and Benjamin, 1987; Bakos, 1991, 1997). Yet, while quality can be judged purely via information available online for intangible products and services such as travel, software or music, tangible products require physical handling and evaluation by the consumer. For example, Clemons, Hann and Hitt (2002) find substantial differences in ticket quality offered by online travel agents, yet these product quality differences can be evaluated online according to criteria such as the number of connections or the accuracy of match between requested and delivered departure/return time.

For tangible products sold in an online channel, Koch and Cebula (2002) point out four categories of products where consumer perception of product quality over the internet is likely to be quite variable: (1) products which involve touch, taste or smell, (2) the sale requires custom fitting, (3) the sales if from a catalog or (4) the sale is accompanied by advice or counsel. This makes intuitive sense since the evaluation of commodity products commonly sold over the Internet such as books, consumer electronics or toys is largely based on brand perception and loyalty since the same product is available at numerous outlets. In contrast, products such as apparel suffer from product

quality perception problems since consumers are unable to physically try on items (Vickery and Agins, 2001). Similarly, grocery items, particularly non-packaged items such as fresh produce, meats and dairy products, are vulnerable to customer misperceptions of product quality because when buying the product online the customer sacrifices the ability to select their own merchandise. Several examinations of Internet grocery retailing point out that customer perceptions of product quality for groceries is likely to be influenced substantially by both the experience level of the customer and the method of picking/assembling the order. First, customers to new to Internet retailing face a hurdle in terms of becoming comfortable with a new way of purchasing a very personal item – after all everyone has different "ideal" characteristics for a piece of fruit such as a pear, thus more experience customers are more likely to rate product quality higher (Tanskanen, Yrjola and Holmstrom, 2002; Ellis, 2003). Second, the method used to pick orders will directly impact quality. Store-based picking is, at best, able to offer comparable or worse actual quality levels – after all, the food has followed all the same steps in the supply chain, including display on store shelves, but the customer has given up control of the process and is trusting the personal shopper to select items for them. In contrast, DC-based picking offers potentially better quality for fresh items because it can cut a link out of the supply chain (instead of producer – DC – store – customer, online ordering allows producer – DC – customer) according to Delaney-Klinger, Boyer and Frohlich (2003).

Based on these arguments, we examine the following hypotheses, which examine the effects of customer experience level and order picking method on product quality perceptions and the relationship between product quality and behavioral intentions:

H3(A): Customer perceptions of product quality will differ by customer experience level.

H3(B): Customer perceptions of product quality will differ based on the picking method for selecting items (i.e. store or DC-based).

H3(C): There is an interaction effect between product experience level and picking method for customer perceptions of service quality.

H4: Product quality is correlated with increased customer satisfaction as measured by behavioral intentions.

Internet Quality

For the purposes of this study, Internet quality measures aspects of quality that are critically affected by taking orders online. In other words, things that customers are likely to view differently online than in a store. In particular, Internet quality includes three components: product range, customer perception of quality vs. that in stores and the rate of substitutions or out of stocks. Let's examine each of these in turn.

The grocery industry in particular and most businesses in general continually grapple with the issue of product range. Offering more variety generally provides customers with more choice, but also makes inventory, replenishment and stocking more difficult. The trade-off between offering more choices and keeping supply chains simple and focused is a classic feature of much of the operations strategy literature (Skinner, 1969; Fisher, 1997). This issue has been particularly challenging for the grocery industry where a typical supermarket stocks 40,000 SKUs, while the typical shopper has about 200 items that account for over 95% of their annual purchases. The vast variety offered by supermarkets comes with inevitable problems with stockouts, over ordering, obsolescence etc. Studies have found consistent problems with stockouts, with 39% of customers unable to find at least one item they are searching for on a typical shopping trip (Emmelhainz, Emmelhainz and Stock, 1991). Thus, the trade-off between range and stockouts is a critical component for grocery retailers (Stassen and Waller, 2002). The offering of groceries for order online is expected to alter the dynamics of this trade-off: companies that choose a DC-based picking approach typically offer fewer SKUs (from 8,000 – 12,000), thus sacrificing some range, while seeking to improve on in-stock availability and reduce substitutions (when an online order is placed

and an item is out of stock, the retailer can skip that item or substitute an alternative item, either alternative is less than perfect for the customer). In contrast, companies that choose a store-based picking approach are likely to have a higher range (typically all or most of the items carried in the store), but have more challenges with availability because of the higher range and because of variability of inventory from store to store.

Our definition of Internet quality includes product range, availability and quality relative to the store. The quality relative to the store could be assumed to be the for Internet ordering, but customers tend not to see it that way because of their unfamiliarity with this channel for ordering groceries (Tanskanen et al., 2002). Thus, we test for differences in Internet quality based on customer experience level and picking method, and also examine the relationship between internet quality and behavioral intentions:

H5(A): Customer perceptions of internet quality will differ by customer experience level.

H5(B): Customer perceptions of internet quality will differ based on the picking method for selecting items (i.e. store or DC-based).

H5(C): There is an interaction effect between product experience level and picking method for customer perceptions of internet quality.

H6: Internet quality is correlated with increased customer satisfaction as measured by behavioral intentions.

Time Savings

One of the most commonly cited reasons for shopping for products on the Internet is to save time. The mass media repeatedly stress the theme that people are continually busy and are looking for ways to save time. Internet ordering is often profiled as one of those potential time savers, since customers can order anytime, anywhere and dressed anyway. Researchers have broadly supported these statements, with Bhatnagar, Misra and Rao (2000) and Donthu and Garcia (1999) both finding that Intenet stores were particularly attractive for time-starved consumers. Researchers such as Eastlick and Feinberg (1999), Bellman, Lohse, and Johnson (1999) and Alreck and Settle (2002)

have examined both customer perceptions of the time savings and potential explanatory factors such as demographics, comfort level with computers and type of shopping. Chen and Hitt (2002) went a step further and examined several factors, including demographics, web site personalization and aspects of the product offering and their relationship with switching and attrition for online brokerages.

In contrast to these studies, our examination of online grocery purchases addresses two factors that have not been addressed in any depth: measurements of the effects of learning or repeat purchases on timesavings and differences in operational execution. Internet ordering certainly can be a time saver, but many people tend to overlook the effects of learning and repeat experience. For example, Amazon is well known for its development and attempted patenting of its one-click ordering system – certainly for repeat customers this offers a significant time savings. However, first time customers of Amazon may actually have to spend more time placing there order due to the need to enter data such as their name, address, billing information etc. – none of which information is required when buying a book in a physical bookstore (Boyer, 2001). Thus, there are likely to be differences between new and repeat online buyers.

The difference between new and repeat purchasers is likely to be more substantial for groceries due to the extreme difference in shopping methods – many people have trouble with online orders for items like cereal since they have a hard time visualizing groceries without handling them. One report by Ellis (2003) indicated that the average time for customers to place their first online order for groceries was 70 minutes, while the average for the fifth order was approximately 30 minutes. Internet ordering of groceries involves a switch in activities from the customer doing their own shopping to a paid employee doing it for them. While the time to place an order online for groceries may seem long, this represents most of the time to order and receive an order when the grocer is picking the order. In contrast, the total time to shop for groceries in a physical store

includes the time to compile a list, travel to the store, shop in the store, checkout and travel home. Yrjola (2001) estimates that the value of customers' time spent shopping represents 20 percent of the value of grocery products, or roughly $90 billion a year in the U.S., given that the annual sales of grocery products are $450 billion per year. In an earlier paper on the online book and CD market, Brynjolfsson and Smith (2000) used a similar method for estimating the time spent shopping in a physical store. However, given the more time intensive nature (more items to select and the need to shop more often) of shopping for groceries and the highly varied opinions of grocery shopping it would be difficult to develop an accurate estimate. Therefore, while estimates such as those used by Brynjolfsson and Smith (2000) and Yrjola provide good insight into the general costs (both financial and time), we believe that customer perceptions are more important since they are the ones must pay for the service.

Based on the above discussion, we examine customer perceptions of time savings, as effected by experience level and picking method. We also examine the relationship between perceived time savings and behavioral intentions:

H7(A): Customer perceptions of time savings will differ by customer experience level.

H7(B): Customer perceptions of time savings will differ based on the picking method for selecting items (i.e. store or DC-based).

H7(C): There is an interaction effect between product experience level and picking method for customer perceptions of time savings.

H8: Time savings is correlated with increased customer satisfaction as measured by behavioral intentions.

Behavioral Intentions

We have already discussed the direct effects of each of the factors on behavioral intentions . We will also examine the moderating effects of customer experience level and picking method on behavioral intentions. Clearly, it is likely that there is a relationship between customer experience

and behavioral intentions, since more experience customers have already expressed their behavioral intentions by making repeat purchases. We also examine the relationship between picking method and behavioral intentions to evaluate if operational differences based on picking method effect behavioral intentions.

H9(A): Behavioral intentions will differ customer experience level.

H9(B): Customer perceptions of behavioral intentions will differ based on the picking method for selecting items (i.e. store or DC-based).

H9(C): There is an interaction effect between product experience level and picking method for behavioral intentions.

METHODS

Sample

The sample consists of customers of five online/home delivery grocers – two in the U.S., two in the U.K and one in Canada. All five firms were generous with their time and allowing us access to their customers, but prefer not to be identified by name given the dynamic nature of the home delivery grocery industry. Thus, we will describe the firms in a general manner while assigning fictional names to each grocer. We are also limited in the degree to which we can describe the individual sales and financial characteristics of these firms due to the highly competitive and developing nature of the industry. We can say that, in aggregate, the firms have annual online/home delivery sales of well over $200 million through over 200 bricks and mortar stores. The customer base of the firms, in aggregate, is well over 200,000 customers, with at least 50,000 loyal or repeat customers that purchase over $500 in groceries per year each. The people in our contact sample account for a total of over 38,000 purchases, and $4.75 million in home delivery sales.

While there are numerous differences, large and small, in the techniques these grocers utilize to take, assemble/pick and deliver grocery orders to customers' homes, we are limited in what can

be revealed by the sensitive nature of the business. Therefore, we will focus on the choice of method to pick grocery orders: either in an existing store or in a distribution center. While there are numerous other operational decisions to be made, this choice is a fundamental strategic decision that forms the foundation for each grocer's operations strategy. Grocers A and B both pick customer orders from existing stores, while grocers C, D and E have all built dedicated distribution centers for picking customer orders.

Data Collection

Table 1 provides a summary of the data collection techniques for the individual grocers in our sample, the number of customers contacted from each company, the number of responses and the response rate. All of the customers contacted had purchased groceries online for home delivery at least once.

Insert Table 1 about here

Our goals for data collection were to receive at least 300 responses for each of the firms in the sample, stratified by the experience level of the customers. In designing the sample of customers to be contacted for each firm, we split each sample into 1/3 brand new customers (those that had placed 1 or 2 orders online from that grocer), 1/3 repeat customers (3 to 6 online orders and 1/3 experienced customers (7 or more online orders). In general, the principles advocated by Dillman's (1978) total design method for survey data collection were followed: initial contact with follow-up reminders, a small incentive for completing the survey and the promise of anonymity in survey responses.

Unfortunately, one of the difficulties involved with working with companies directly to contact their customers is that each company wants data collection handled in a separate manner. In

general, companies are leery of allowing outsiders to contact customers due to recent publicity about revealing customer sensitive information. Furthermore, companies offering online grocery ordering are extra sensitive due to the dual need to build product awareness/trust and the desire to avoid being classified as "another internet startup" like Webvan. We thus had to negotiate with each individual company to balance their desires to protect their customers from undue spamming with our desire to employ identical methods across multiple companies. As can be seen in Table 1, we were able to employ substantially similar data collection methods albeit with some minor exceptions. There are five basic data collection methods that differed across grocers: sample selection, invitation to customer, incentive offered, survey method and follow-up invitation.

Ideally every company would have a stratified contact sample with equal numbers of new (1 or 2 orders), repeat (3 to 6 orders) and experienced (7 or more) customers. In addition to facilitating comparison of customers based on experience level (i.e. number of purchases), stratified samples also allow us to track specific customer responses so that post hoc, longitudinal information can be gathered at a later date. This was done in 3 out of 5 cases by the grocers contacting the customer directly and referring them to our independent website. There, customers were asked to input an ID number that would allow us to match up their past/future purchasing history without knowing their personal information (i.e. name, address or email address). One of 5 companies (Grocer B) was unwilling to identify customers in any manner (even using a single-blind approach as described above where only personnel at the grocer would be able to identify specific customers). Thus, grocer B employed an opt-in approach where customers were invited to participate in the survey at checkout for their order. This method resulted in a randomized sample, since all customers received the invitation to participate. We added a question on this survey asking customers how many times they had shopped with grocer B, thus we were able to stratify this sample in a post hoc manner based on customer responses.

Another minor methodological difference can be seen with grocer D. This grocer did not want to be perceived as an internet startup, preferring to approach customers as a grocery company that happened to take orders over the internet. Therefore, grocer D asked us to send a written letter to customers with a written survey. The standard techniques of two follow-up letters and a pre-paid, business reply envelope (Dillman, 1978) were employed for this survey. It is interesting that this technique resulted in by far the highest response rate (74.6% for grocer D vs. 34.2% for the next highest company).

The two final methodological differences relate to the offering of a small incentive for survey completion and follow-up invitations to participate in the study. Four of the five grocers provided a small incentive to customers to participate in the study. These incentives ranged from a company hat for all customers filling out a survey, to one free delivery, to a raffle for 20 pairs of movie theater tickets. In all, the incentives were all worth less than $10 per customer, but these companies felt that it was important to compensate their busy customers for their time and valuable feedback. With regard to follow-up invitations, three of the five grocers sent a reminder (either by email or by written mail) to customers one week after the initial invitation to participate in the study. Grocer B was unwilling to identify customers in any manner or to directly contact customers, thus there was no follow-up at all. In contrast, Grocer E sent the initial invitation to 2000 customers via email, but felt that they did not want to risk upsetting customers with another reminder.

As shown in Table 1, the overall response rate for the entire sample is 2,985 responses out of 16,577 customers contacted, or 18.0%. All of the data was collected in the period August, 2002 – May, 2003. This compares very favorably to the response rate in similar studies (Duray, Ward, Milligan and Berry, 2000; Papke-Shields, Malhotra and Grover, 2002). With the exception of grocer B, all of the individual company response rates were well above 20%. Grocer B had a

substantially lower response rate due to the different data collection methodologies employed (opt-in rather than special invite, no incentive and no follow-up invitation). Excluding grocer B from the sample, the overall response rate is 31.2%. To assess non-response bias, we conducted chi-square tests on the proportion of positive responses for the number of orders placed online with the sponsoring grocer. Grocers A and D indicated no potential for bias, whereas grocers C and E had significant chi-square statistics. The data indicate that customers that have placed more orders with a grocer were more likely to complete the survey, a result which is both intuitively logical and has been observed in the literature. This finding is consistent with our primary reason for stratifying our contact samples by the number of orders placed: to examine customer differences based on usage. Therefore, our analysis of the data will take this into account and perform separate analyses on fairly new (1 – 4 orders) and fairly experienced (7 or more orders) users.

Several tests were made across the different sub-samples to test for potential biases due to the different data collection methods (physical mail, email or Internet) and none of the tests suggested the presence of a bias. This result is consistent with the findings of Couper (2000) and Klassen and Jacobs (2001) that surveys can be administered via physical mail, electronic mail or the Internet with no cause for concern as long as the research design is solid and the questionnaire is consistent.

Scales

This section describes the scales used to measure the various components of Figure 1. We used existing scales where possible, but also tried to develop customized scales where appropriate to capture the dynamic and customized nature of the online grocery shopping. Each of the factors shown in Figure 1 is described below, while the following section will describe the measurement analysis. The individual items included in each scale are shown in the Appendix.

The construct of service quality has been studied and debated for the last decade (e.g., Cronin and Taylor 1992; Teas 1993; Van Dyke, Prybutok, and Kappelman 1999). In general, recent evidence supports the use of performance perceptions in measures of service quality (Zeithaml, Berry, and Parasuraman 1996). Because of the need to ensure construct and measurement equivalence across multiple grocers in multiple countries, it is especially important to use a broad range of scale items that can be generalizable across the grocers and countries while at the same time keeping the items to a manageable number. As such, we devised a scale composed of ten items based on Parasuraman, Zeithaml, and Berry's (1985) ten original dimensions of service quality. Similar scales have been used by Kettinger, Lee and Lee (1995).

The product quality scale is a new scale that we developed to assess customer perceptions of product quality relative to what they can get in the store. This scale measures cusomters' view of the general brand quality of the physical products. In contrast, internet quality measures aspects of operational execution – i.e. how well the grocer delivers the product. Thus, this scale includes items relating to range of product choices and the number of substitutions for out of stock items. The third item included in Intqual relates to quality of products relative to that obtained in the store. These three issues repeatedly came up during our interviews with managers at all of the grocers in our study as key concerns of customers as well as key challenges for the company to execute well. Time savings is composed of two items: one asking whether the time to place an order shortens with repeat experience and one asking customers to rate the total time to place and receive an order online relative to traveling to the store and back. It is interesting to note that the average response across all five grocers for this item (b) was 5.98, suggesting that customers find online ordering to be substantially better than traditional shopping. The perceptual measure of time savings correlates strongly with customers objective measure (minutes to place order) with a correlation of -0.33 ($p < 0.01$). We use the perceptual measure of time savings for two reasons: first, because it measures

how customers view and the time savings and second, because the responses for the likert scale questions are more complete (the objective question is missing on a large percentage of surveys).

The indicators of behavioral intentions represent the outcome measures in this study. Zeithaml, Berry, and Parasuraman (1996) suggest that positive behavioral intentions are reflected in the service provider's ability to get its customers to: (a & b) remain loyal to them, (c) pay price premiums, (d) communicate concerns to other customers and (e) communicate concerns to the company.

ANALYSIS AND RESULTS

Measurement Analysis

Prior to hypothesis testing, the multi-attribute measures were assessed with respect to (1) reliability and validity and (2) potential common method variance issues.

Reliability and Validity. Table 2 reports the correlations among the study measures. Table 3 presents the basic statistics (i.e., means and standard deviations) and the results of the measurement analysis (e.g., average variances extracted, construct reliabilities, loadings, and fit indices). Overall, the five latent constructs, involving 23 items, were found to be reliable and valid in the context of this study. The details of the measurement analyses are discussed in the remainder of this section.

Insert Tables 2 and 3 about here

Following data collection, we assessed scales' dimensionality, reliability, and validity. First, the psychometric properties were evaluated in one confirmatory factor analysis (CFA) using LISREL 8.54 (Jöreskog, Sörbom, Du Toit and Du Toit, 2000). Next, unidimensionality and discriminant validity of the constructs were assessed by examining the average variance extracted

for each scale and comparing it with the shared variances of all scale combinations (Fornell and Larcker,1981

The model fit was evaluated using the DELTA2 index (Bollen, 1989), the relative noncentrality index (RNI)(McDonald and Marsh, 1990), and the comparative fit index (CFI) (Bentler, 1990), which have been shown to be the most stable fit indices by Gerbing and Anderson (1992). After removing one trouble item (item #4 of behavioral intentions), the CFA model resulted in a good fit to the data, with DELTA2, RNI, and CFI all being .97 for the overall model (see Table 3; χ^2=4114.95, df=199). This analysis shows that the CFA model is a good fit to the data; thus, we continued analyzing the individual measures and their indicators.

Within the confirmatory factor analysis setting, composite reliability was calculated using procedures outlined by Fornell and Larcker (1981). The formula specifies that: $CR_\eta = (\Sigma\lambda\gamma_i)^2 / [(\Sigma \lambda\gamma_i)^2+(\Sigma\varepsilon_i)]$, where CR_η = composite reliability for scale η, λ_{yi} = standardized loading for scale item γ_i, and ε_i = measurement error for scale item γ_i. We also examined the parameter estimates and their associated t-values, and assessed the average variance extracted for each construct (Anderson and Gerbing, 1988; Bagozzi and Yi, 1988). Average variance extracted was calculated using the following formula: $V_\eta = \Sigma\lambda\gamma_i^2 / (\Sigma\lambda\gamma_i^2+\Sigma\varepsilon_i)$, where V_η = average variance extracted for η , λ_{y_i} = standardized loading for scale item γ_i, and ε_i = measurement error for scale item γ_i. The composite reliabilities for the seven constructs ranged from .55 to .96; the factor loadings ranged from .30 to .91 ($p < .01$), and the average variances extracted ranged from 34.50% to 70.70% (Table 2).

In assessing discriminant validity, we followed the procedure recommended by Fornell and Larcker (1981). This procedure entails comparing a construct's average variance extracted (AVE) with the shared variances (SV) between that construct and all other scales. In this analysis, we

found that the AVE was higher for each construct than the corresponding SVs (see Tables 1 and 2). Specifically, in the overall sample, the AVE for the service quality scale is 70.70% and the SVs range from 4% to 19%; the AVE for product quality is 69.33% and the SVs range from 6% to 35%; the AVE for internet quality is 46.00% and the SVs range from 12% to 35%; the AVE for time savings is 34.50% and the SVs range from 4% to 13%; and the AVE for behavioral intentions is 41.50% and the SVs range from 13% to 20%. Thus, overall, the measurement analysis indicated that the five constructs and their purified 23 scale items were reliable and valid in the context of this study.

Common Method Variance. We used Harmon's One-Factor test to examine potential common method variance issues. In each of these sample scenarios, the 23 items were factor analyzed using SPSS 11.0, with a principal component extraction and varimax rotation, to examine if one single factor would emerge and/or if one general factor would account for most of the covariance in the variables (e.g., Podsakoff and Organ, 1986). In this analysis, we found that the items loaded on their conceptually predetermined factor and achieved variances explained ranging from a high of 33.01% to a low of 5.28% for the five scales (for a total explained variance of 70.31%). Thus, common method bias does not appear to be an inhibiting factor in the hypothesis testing. As such, we form our final scales by taking the average of the items in each scale and proceed to by examining the research hypotheses.

Results of Hypothesis Testing

Our results follow the general outline shown in Figure 1: the odd hypotheses (H1, H3, H5 and H7) regarding the moderating factors of customer experience level and picking method are tested by means of ANOVA for both main and interaction effects. The second step is to test the relationship between each of our direct factors (Service Quality, Product Quality, Internet Quality

and Time Savings) and outcomes (Behavioral Intentions) using multiple regression for each of the four combinations of customer experience level and picking method.

Moderating Factors

Table 4 show the ANOVA results when Customer Group (New or Repeat) and Pick Method (DC-based or store-based) are used as the independent variables. Table 4 shows the means for the overall sample as well as each of the four combinations of Customer and Pick Groups, as well as the F-statistics and significance values for the main effects and interactions effects.

Insert Table 4 about here

A quick review of Table 4 reveals that Customer Group has a significant effect for 4 of the 5 dependent variables, Pick Group is significant for two out of five and that there is an interaction effect for four out of five. Some interesting insights are found by examining each independent variable. Service Quality is rated significantly higher by Repeat Customers than New Customers, while there is no effect for Pick Method and no interaction effect. This suggests that more experienced customers value and appreciate the service available from home delivery as they use the service more. It is not surprising that there is no effect for Pick Method as service quality primarily relates to interactions between customers of the company and workers, since workers picking orders are primarily back-office and invisible to customers the choice of picking method is unlikely to affect service quality.

The results for Product Quality show that there is both a main effect for Customer Group and an interaction effect. There is a small difference between New (mean = 5.27) and Repeat Customers (mean = 5.41). This result indicates that more experienced customers rate Product Quality higher, possibly because they become more accepting of someone else selecting their groceries for them. The interaction effect can be clearly seen in Figure 2, which shows that the

ratings of Repeat Customers increase substantially over New Customers when groceries are picked from a DC. In contrast, the ratings of Product Quality decrease for Repeat customers. This finding supports the hypothesis that delivering groceries straight from a DC can provide better quality, but that customers may take some time to get used to this idea.

Insert Figures 2 - 5 about here

The results for Internet Quality show no main effect for Customer Group, but significant effects for the Pick Group and the interaction. Figure 3 graphically shows these effects. Store Picking outperforms DC-based picking at both levels of Customer Experience, but the gap between Store and DC-Pick is much smaller for Repeat Customers. This raises the question of what the trend is with this finding – will this gap continue to close as customers gain more experience. Similarly, what effect does the company's choice of product range and inventory level have? We will discuss these further in the next section.

Time Savings has significant main effects for both Customer Group and Pick Group, as well as an interaction effect. As shown in Figure 4, New Customers do not perceive much difference between the two pick methods, but Repeat Customers find a much-higher Time Savings from the DC-Pick Method versus the Store-based Pick Method. The data for Behavioral Intentions shown in Table 4 and Figure 5 exhibit a very similar pattern. There is a strong main effect for Customer Group as well as an interaction effect. As shown in Figure 5, The DC-Pick Method has much higher ratings for Repeat Customers than the Store-based Pick Method.

To summarize, our data provide support for Hypotheses 1(A), 3 (A and C), 5 (B and C), 7 (A, B and C) and 9 (A and C). The data indicate that both customer experience level and picking method have strong effects on customer perceptions of three aspects of quality, time savings and behavioral intentions.

Multiple Regression Results – Predicting Behavioral Intentions

We test the relationships between the direct factors shown in Figure 1 and behavioral intentions in two steps. We first test the overall model for the entire sample, then we test individual models for each of the four combinations of Customer Group and Picking Method.

Table 5 shows the regression equation derived for the entire sample. We show the Beta weights for the un-standardized data because this allows greater interpretability. The combination of Service Quality, Product Quality, Internet Quality and Time Savings provides an R^2 value of 0.34. All four of the independent variables are significant at the $p < 0.01$ level. Based on the magnitude of the coefficient, Service Quality has the largest effect, followed by Time Savings. This is interesting since increasing Service Quality can often be considered as directly conflicting with Time Savings in an operational setting, since better Service Quality often requires spending more time with customers. It is our belief that grocery home delivery offers an opportunity to capitalize on a clear moment of truth at the time of delivery wherein high service quality can be offered in a time efficient manner. We will discuss this possibility in the next section. For now however, the results in Table 5 indicate that Behavioral Intentions can be predicted with a high degree of accuracy and that all four of the independent variables make a contribution.

Insert Tables 5 and 6 about here

Tables 6A and 6B show separate regression results for Store-based Picking and DC-based Picking, respectively, and for New and Repeat Customers within each of these groups. First, all four R^2 values exceed those for the overall sample, which supports the proposition that Customer Group and Picking Method moderate customer perceptions. Second, comparing the beta weights for the different regressions provides useful insights. In table 6A, New Customers do not have a significant relationship between Internet Quality and Behavioral Intentions, furthermore the beta

coefficient (0.11) is the lowest of the four independent variables. Similarly, Time Savings has a less significant relationship ($p < 0.05$) and a relatively lower coefficient. This may be due, in part, to the smaller sample size for this group, but it may also be because New Customers do not recognize the trade-offs inherent in home delivered groceries – namely that they are not in control of making substitutions for out-of-stock items. In contrast, all four variables are significant for Repeat Customers. It is interesting that the beta coefficient increases for Internet Quality (0.18 versus 0.11 for New Customers) while the coefficient for Product Quality decreases (0.12 for Repat Customers versus 0.25 for New). These findings imply that Repeat Customers place more weight on Internet Quality as they place more orders and less weight on Product Quality. This may be because more experienced customers discover the drawbacks of limited range and product substitutions, and thus learn to value good execution with respect to these facets. In contrast, more experienced customers may be more comfortable accepting the quality of products that they do not personally select once they have used the service a few times and gained confidence through a series of successful orders.

Table 6B shows regression equations for DC-based picking for both New and Repeat Customers. Both R^2 values are quite high (0.34 and 0.37) and all four independent variables are significant. In comparison to the beta coefficients for Store-based Picking (Table 6A), the coefficients for DC-based picking are much more homogenous. The coefficients for New Customers only vary between 0.19 and 0.24, while the coefficients for Repeat customers vary between 0.19 and 0.28. In addition, the coefficients do not change much when moving from New to Repeat Customers. This implies two things. First, the four factors (Service Quality, Product Quality, Internet Quality and Time Savings) are all of roughly comparable importance for keeping customers of DC-based Picking satisfied. Second, customer perceptions do not change substantially as they gain experience – in other words, companies that pick from a DC must perform equally well on all four dimensions – they do not get a "free pass" in any one area once customers reach a certain

level of experience. This may place great pressure on these companies since any mistake is likely to have severe repercussions, in contrast to Store-based Picking where customers may be more forgiving of something like Product Quality errors because of their prior experience with the bricks-and-mortar arm of the grocer.

To summarize, the data shown in Tables 5 and 6 provide support for the hypotheses H2, H4, H6 and H8 regarding the relationship between the direct factors shown in Figure 1 and behavioral intentions of customers. In both the overall model and the individual combinations of customer level and pick group, the behavioral intentions of customers can be predicted quite accurately. However, the relative weight of each independent factor differs substantially based on customer and pick group. We turn now to a discussion of the implications of these findings for grocery home delivery companies.

DISCUSSION

In examining home delivered groceries, our evidence supports the twin propositions that customer experience level and order-picking method have an effect on several different dimensions of the overall customer experience. In general, customers with more experience rate the experience more highly – in terms of service quality and product quality, time-savings and behavioral intentions. This finding provides support for the idea that customers must be re-trained to accept a new method of shopping for groceries, after all they have been shopping in-person and in the store for their entire lives so it is logical that placing orders online for home delivery, while convenient, may offer some initial hurdles to adoption. Our second major finding is that the method for picking customer orders has a significant effect on product and internet quality, time-savings and behavioral intentions – either as a main effect or as an interaction with customer experience level. This finding supports the hypothesis that picking from a distribution center can actually provide better product

quality because the shortened supply chain, but that customers may take a while to realize advantages because of their concern about dealing with a far off DC that they can not physically see or touch.

As shown in Figures 2 and 3, the interaction effects for Product Quality and Internet Quality are quite substantial. From an observer's point of view, it is logical that picking orders from a DC offers advantages due to a shortened supply chain – a la Dell computers or Amazon books. In comparison, picking orders from existing stores simply adds another link to the supply chain – rather than customers selecting their own groceries, the store pays an employee to select orders and then deliver them. However, there are also numerous challenges in picking orders from DCs, including convincing customers of the utility of this approach and mastering the numerous intricacies of piece picking orders in high volume for a wide variety of items. Clearly, early competitors such as Webvan or HomeGrocer were not able to master these challenges before running out of cash. In contrast, the second generation of DC-based grocers (Ocado, FreshDirect, Grocery Gateway, SimonDelivers) shows some signs of mastering the learning curve. Our data generally indicate that DC-based picking outperforms store-based picking on several measures of customer perception.

In examining the ability to predict customer behavioral intentions, all four of the scales assessed in this study have a strong effect. Home delivery grocers must simultaneously deliver service, product and internet quality as well as a substantial time-savings. The relative importance of each of these four factors differs substantially for new vs. repeat customers and store-based vs. DC-based picking. The data indicate that established customers (repeat) of DC-based picking grocers are quite demanding in terms of all four factors. This finding is logical given the new format of shopping and the inability of shoppers to visit their neighborhood store for these retailers. In fact, this is a theme that the home delivery grocers have adopted for their marketing – Ocado in

the U.K. asks their customers to “Be more Demanding” and quit tolerating many of the common hassles of traditional, in-store grocery shopping. In closing, we argue that home delivered groceries are helping to re-shape the expectations of customers, who are slowly becoming less tolerant of the many inconveniences and willing to pay a slight premium for better service. Furthermore, DC-based grocers are showing that there is potential in this business model. Future research should continue to track developments in this nascent branch of the ubiquitous grocery industry. In particular, researchers should continue to examine three areas: how customer perceptions evolve, how retailers manage the receipt and picking of orders, and how deliveries to the final customer are best managed.

Table 1. Description of Participating Companies and Data Collection Methods

		Grocer A	Grocer B	Grocer C	Grocer D	Grocer E
Data Collection Methods						
Sample Selection		Stratified	Random	Stratified	Stratified	Stratified
Invitation to Customer		Email	Opt-in at checkout	Email	Written letter	Email
Incentive		Yes	No	Yes	Yes	Yes
Survey Method		Web Survey	Web Survey	Web Survey	Written Survey	Web Survey
Follow-up invitation		Yes	No	Yes	Yes	No
Customers Contacted	16,577	1,159	10,418	2,500	500	2,000
Responses	2,985	396	1,066	690	373	460
Response Rate	18.0%	34.2%	8.6%	27.6%	74.6%	23.0%
Pick Method		Store	Store	DC	DC	DC

Table 2. Correlation Matrix

	Service Quality	Product Quality	Internet Quality	Time Savings	Behavioral Intentions
Service Quality	1.00	0.19	0.15	0.04	0.19
Product Quality	0.44	1.00	0.35	0.06	0.20
Internet Quality	0.39	0.59	1.00	0.12	0.20
Time Savings	0.21	0.25	0.35	1.00	0.13
Behavioral Intentions	0.44	0.45	0.45	0.36	1.00

Note: All correlations are significant at the $p < 0.01$ level.
Correlations are included below the diagonal.
Shared variances are included above the diagonal.

Table 3. Means, Standard Deviations, and Measurement Statistics

Variable	Mean	Standard Deviation	Variance Extracted	Composite Reliability	Factor Loadings
Service Quality	6.11	0.91	70.70%	0.96	0.75 to 0.90
Product Quality	5.40	1.09	69.33%	0.87	0.78 to 0.91
Internet Quality	5.12	1.22	46.00%	0.72	0.65 to 0.71
Time Savings	5.74	1.05	34.50%	0.55	0.30 to 0.78
Behavioral Intentions	5.30	1.01	41.50%	0.73	0.39 to 0.77

Fit Statistics

χ^2	4114.95
d.f	199
DELTA2	.97
RNI	.97
CFI	.97

Table 4. ANOVA Comparison for Customer Type (New or Repeat) and Order-Picking method (DC or store)

Customer Group		New	Repeat	New	Repeat	F-Statistics		
Pick Group		Store-based	Store-based	DC-based	DC-based	Customer		
OVERALL SAMPLE						Group	Pick Group	Interaction
Service Quality								
Mean	6.06	5.94	6.08	6.02	6.15	9.51**	2.59	0.01
St. Dev.	0.87	0.93	0.92	0.83	0.84			
Product Quality								
Mean	5.35	5.40	5.35	5.24	5.51	4.20*	0.00	8.22**
St. Dev.	1.10	1.20	1.18	1.07	0.95			
Internet Quality								
Mean	5.20	5.46	5.39	4.99	5.18	0.89	33.58**	5.00**
St. Dev.	1.15	1.10	1.23	1.11	1.07			
Time Savings								
Mean	5.74	5.61	5.76	5.59	5.99	26.45**	4.21*	5.84*
St. Dev.	1.03	1.21	1.04	1.00	0.91			
Behavioral Intentions								
Mean	5.22	5.02	5.35	4.95	5.55	87.06**	1.85	7.13**
St. Dev.	1.01	1.07	0.98	1.00	0.89			
	N = 2150 Overall	N = 179	N = 699	N = 795	N = 472			

NOTES:
* $p < 0.05$
** $p < 0.01$

Figure 1. Research Model

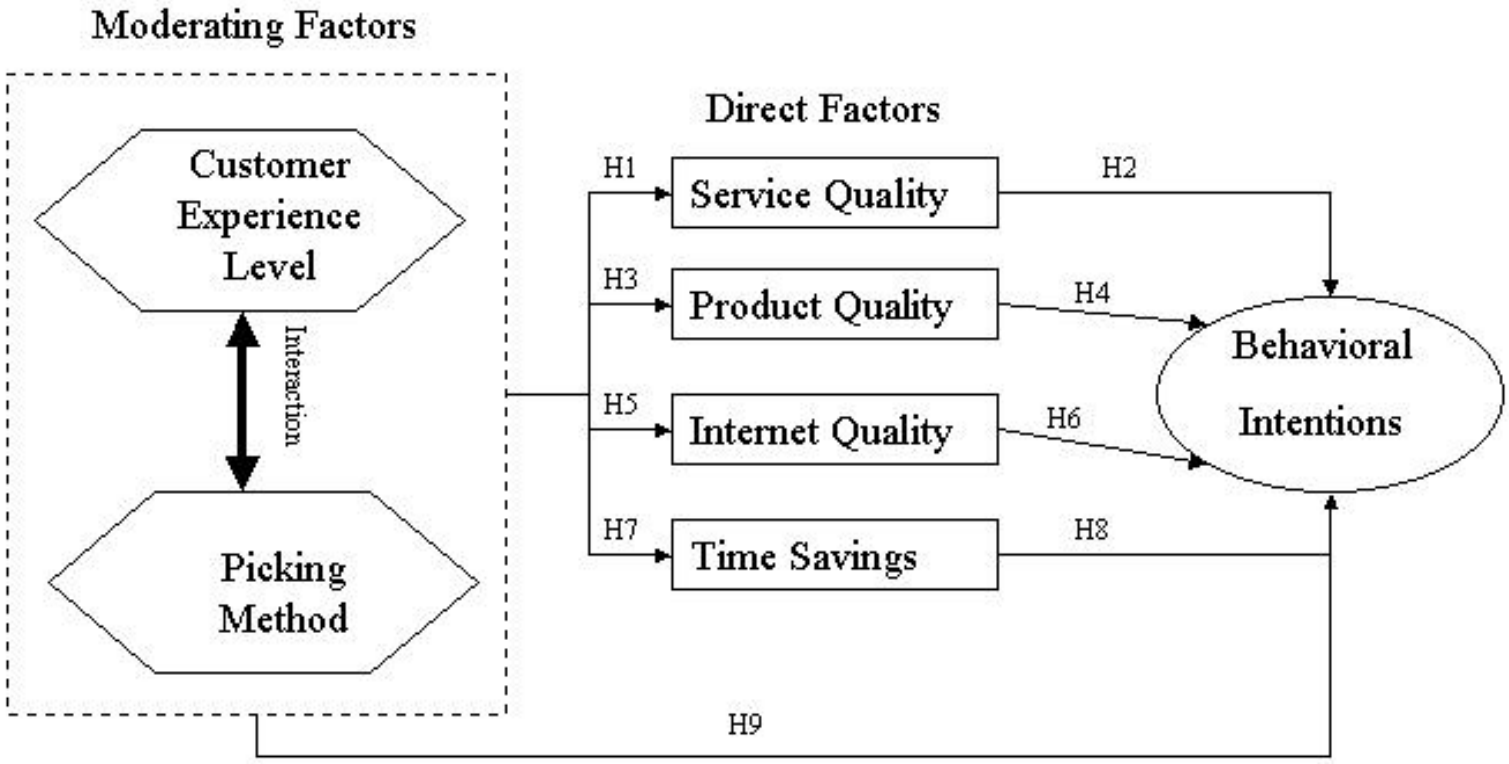

Figure 2. Interaction Effect for Product Quality

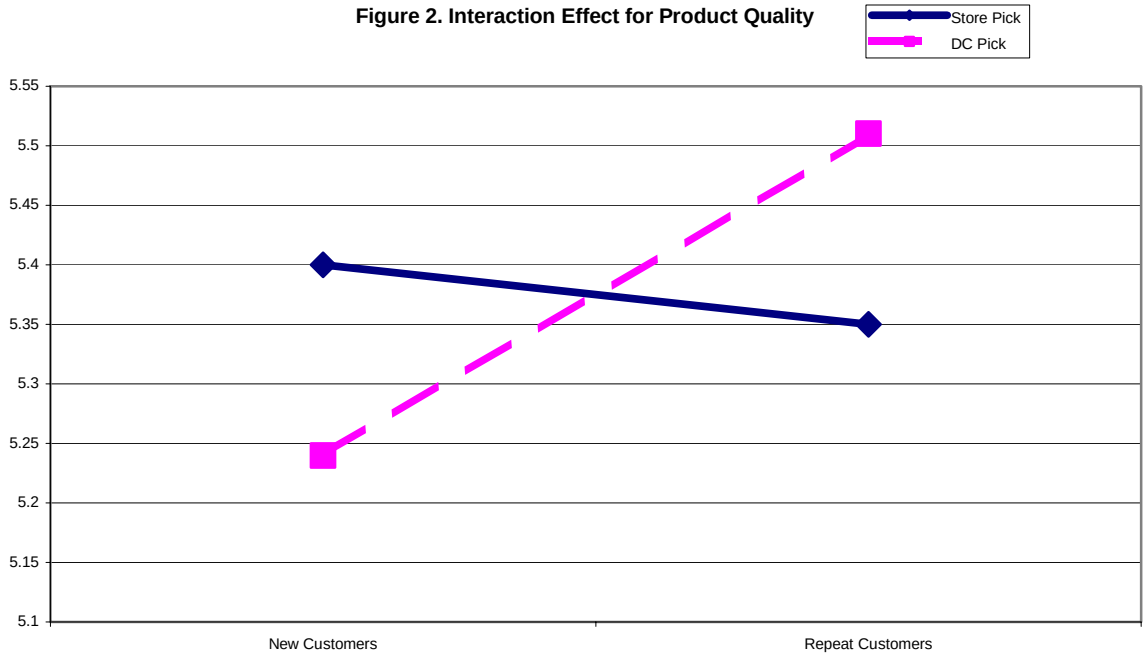

Figure 3. Interaction Effect: Internet Quality

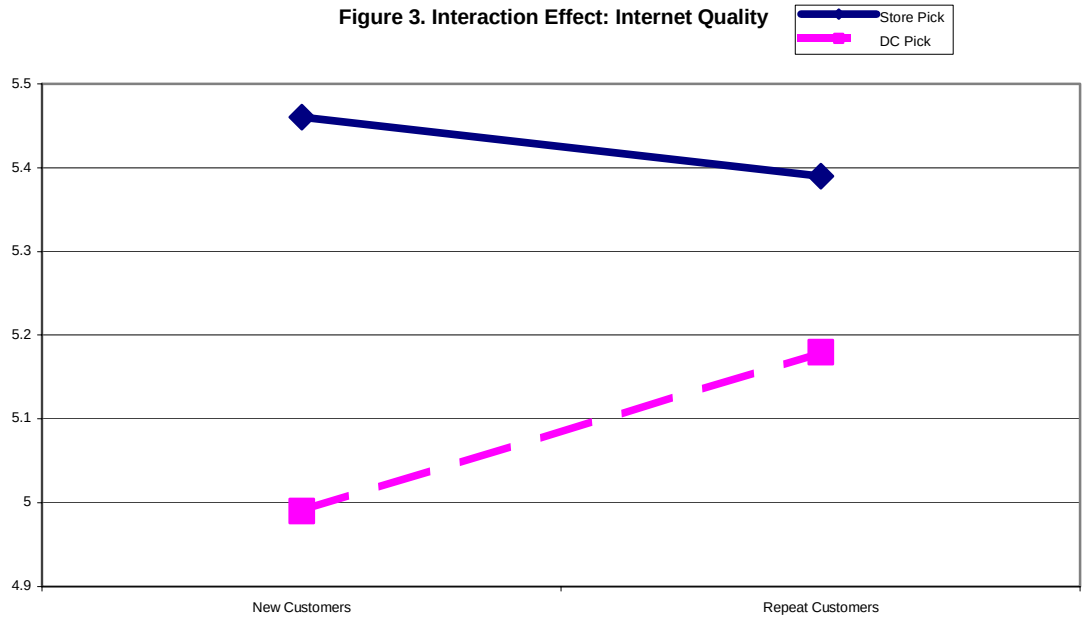

Figure 4. Interaction Effect: Time Savings

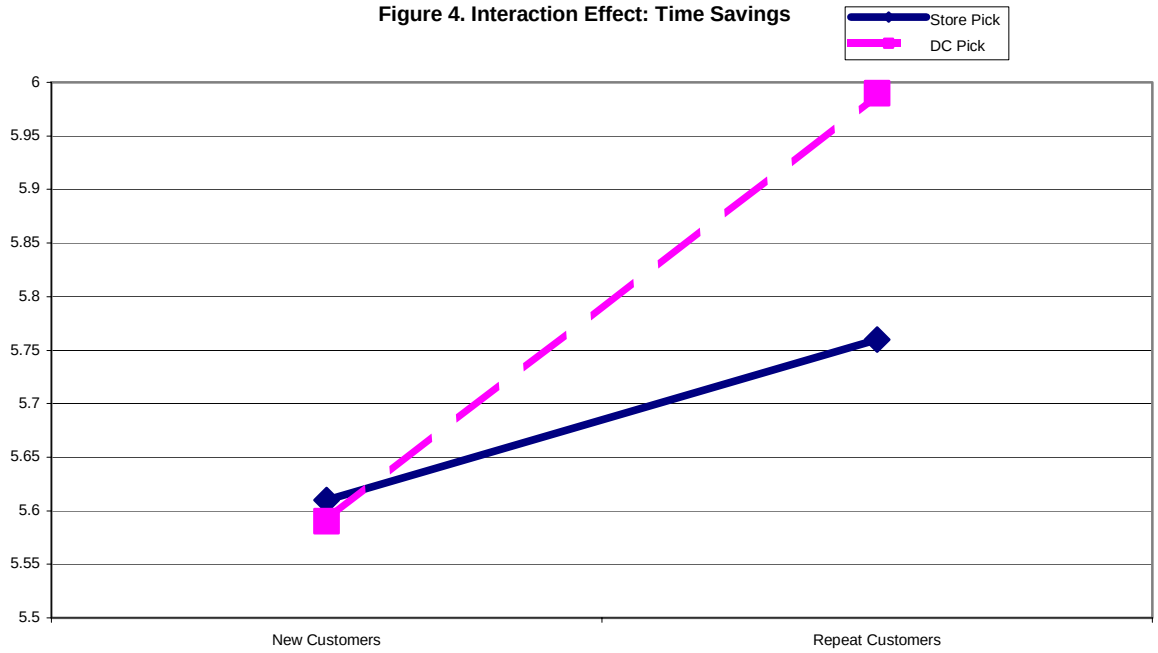

Figure 5. Interaction Effect: Behavioral Intentions

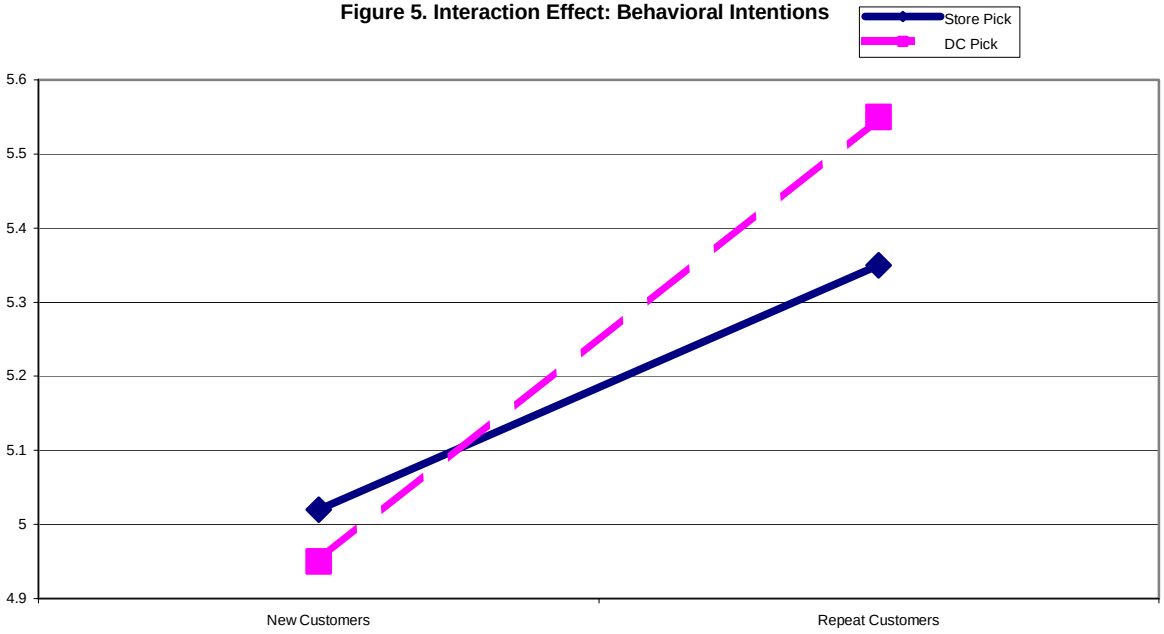

Table 5. Linear Regression with Behavioral Intentions as Dependent Variable

Variable	Beta
Constant	0.76**
Service Quality	0.25**
Product Quality	0.18**
Internet Quality	0.17**
Time Savings	0.20**
R^2	0.34**
N = 2150	

NOTES:
* $p < 0.05$
** $p < 0.01$

Table 6. Linear Regression by Pick and Customer Type with Behavioral Intentions as Dependent Variable

A. Regressions for Store-based Picking

	New Customers (1 – 4 orders)	Repeat Customers (7 or more orders)
	Beta	Beta
Constant	0.69**	1.60**
Service Quality	0.29**	0.22**
Product Quality	0.25**	0.12**
Internet Quality	0.11	0.18**
Time Savings	0.13*	0.14**
R^2	0.37**	0.34**
	N = 177	N = 698

B. Regressions for DC-based Picking

	New Customers (1 – 4 orders)	Repeat Customers (7 or more orders)
Variable	Beta	Beta
Constant	0.40**	0.78**
Service Quality	0.24**	0.28**
Product Quality	0.19**	0.18**
Internet Quality	0.20**	0.18**
Time Savings	0.20**	0.19**
R^2	0.34**	0.37**
	N = 793	N = 472

NOTES:
* $p < 0.05$
** $p < 0.01$

Appendix. Scales

All Scales rated from 1 = strongly disagree to 7 = strongly agree except where noted.

SERVICE QUALITY

a. Grocer X employees are reliable in providing the service I expect
b. Grocer X employees are understanding of my service needs
c. Grocer X employees are responsive to my service requests
d. Grocer X employees are competent in providing the expected service
e. I feel secure in service encounters with Grocer X employees
f. Grocer X employees are courteous in providing me service
g. Grocer X employees are available to answer my service-related questions
h. The tangible (appearance of trucks, staff, products) aspects of
i. Grocer X service are excellent
j. Grocer X has good credibility in providing the service I need
k. I have access to communicate with Grocer X regarding my service needs

PRODUCT QUALITY

a. Grocer X has prestigious products
b. Grocer X has an excellent assortment of products
c. Grocer X products are among the best

INTERNET QUALITY

a. GrocerX has a sufficient range of product choices (I can get what I want)
b. The products are the same quality as I can get in the store
c. The number of substitutions or out of stock items is reasonable

TIME SAVINGS (inter-item correlation = 0.20, $p < 0.01$)

a. The time to place an order becomes much shorter as I use the system more
b. The total time (including travel time to and from the store) to place an order and pickup groceries at the store

Item b asked customers to rate their experience with Online ordering vs. their experience with a "traditional" store. The specific question asked was:

> Please rate the degree of change when using the Internet for ordering groceries through Grocer X in comparison to shopping in a neighborhood store (i.e. an Albertsons, Publix, Kroger, Safeway etc.). The question was rated on a scale from 1 = Much worse than in-store shopping, to 4 = about the same to 7 = Much better than in-store shopping.

BEHAVIORAL INTENTIONS

a. I would classify myself as a loyal customer of Grocer X.
b. I do not expect to switch to another online grocer to get better service in the future
c. I would continue to do business with Grocer X even if I had to pay more
d. I would complain to other customers if I experienced a problem with Grocer X service*
e. I would complain to Grocer X employees if I experienced a problem with their service

NOTES:

* Indicates this item was dropped from the scale based on analysis of inter-item variability.

Insert name of specific company wherever Grocer X appears.

REFERENCES

Alreck, P.L. and Settle, R.B., (2002), "The Hurried Consumer: Time-saving Perceptions of Internet and Catalogue Shopping", *Journal of Database Marketing*, 10(1), 25-35.

Andersen Consulting, "The Daunting Dilemma of Trade Promotion", 1997.

Anderson, J.C. and Gerbing, D.W. (1988), "Some Methods for Respecifying Measurement Models to Obtain Unidimensional Construct Measurement," *Journal of Marketing Research*, 19 (November), 453-460.

Bagozzi, R.P. and Yi, Y. (1988), "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16 (2), 74-94.

Bakos, J. Y., (1991) "A Strategic Analysis of Electronic Marketplaces", *MIS Quarterly*, 15 (3), 295–310.

Bakos, J. Y (1997), "Reducing Buyer Search Costs: Implications for Electronic Marketplaces", *Management Science*, 43(12), 1676-1692.

Bellman, S., Lohse, G.L. and Johnson, E.J., (1999), "Predictors of Online Buying Behavior", *Communications of the Association for Computing Machinery*, 42(12), 32-38.

Bhatnagar, A., Misra, S. and Rao, H.R., (2000), "On Risk, Convenience and Internet Shopping Behavior", *Communications of the Association for Computing Machinery*, 43(11), 98-105.

Bitner, M.J., (1992), "Servicescapes: The Impact of Physical Surroundings on Customers and Employees", *Journal of Marketing*, 56(2), 57-701.

Bollen, Kenneth A. (1989), *Structural Equations with Latent Variables*, New York: Wiley & Sons.

Boulding, W., Kalra, A. and Staelin, R., (1999), "The Quality Double Whammy," *Marketing Science*, 18 (4), 463-484.

Boyer, K.K., (2001), "E-Operations: A Guide to How the Internet Streamlines Operations", *Business Horizons*, 44(1), 47-54.

Brynjolfsson, E. and Smith, M.D., (2000), "Frictionless Commerce? A Comparison of Internet and Conventional Retailers", *Management Science*, 46(4), 563-585.

Chen, P.Y. and Hitt, L.M., (2002), "Measuring Switching Costs and the Determinants of Customer Retention in Internet Enabled Businesses: A Study of the Online Brokerage Industry", *Information Systems Research*, 13(3), 255-274.

Clemons, E. K., Hann, I. H., and Hitt, L. M., (2002) "Price Dispersion and Differentiation in Online Travel: An Empirical Investigation", *Management Science*, 48 (4), 534 – 549.

Couper, M. P. (2000), "Usability Evaluation of Computer-Assisted Survey Instruments", *Social Science Computer Review*, 18 (4), 384-396.

Cronin, J.J. and Taylor, S.A., (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56 (3), 55-68.

Delaney-Klinger, K., Boyer K.K. and Frohlich, M., (2003), "The Return of Online Grocery Shopping: A Comparative Analysis of Webvan and Tesco's Operational Methods", *The TQM Magazine*, 15 (3), 187-196.

Dillman, D. A. (1978), *Mail and Telephone Surveys: The Total Design Method*, Wiley, New York, NY.

Donthu, N. and Garcia, A., (1999), "The Internet Shopper", *Journal of Advertising Research*, 39(3), 52-58.

Duray, R, Ward, P, Milligan, G., and Berry, W., (2000), "Approaches to Mass Customization: Configurations and Empirical Validation", *Journal of Operations Management*, 18 (6), 605-626.

Eastlick, M.A. and Feinberg, R.A., (1999), "Shopping Motives for Mail Catalog Shopping", *Journal of Business Research*, 45(3), 281-290.

Ellis, C., (2003), "Lessons from Online Groceries", *Sloan Management Review*, 44 (2), 8.
Emmelhainz, L.W., Emmelhainz, M.A. and Stock, J.R., (1991), "Logistics Implications of Retail Stockouts", *Journal of Business Logistics*, 12(2), 129-142.
Fisher, M., (1997), "What is the Right Supply Chain for Your Product", *Harvard Business Review*, 75 (2), 105-116.
Fornell, C. and Larcker, D.F., (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39-50.
Frankel, R., Goldsby, T.J. and Whipple, J.M., (2002), "Grocery Industry Collaboration in the Wake of ECR",**.** *International Journal of Logistics Management*, 13(1); 57-72.
Gerbing, D.W. and Anderson, J.C., (1992), "Monte Carlo Evaluations of Goodness of Fit Indices for Structural Equation Models," *Sociological Methods and Research*, 21 (November), 132-160.
Hamilton, A., (2003), "What's for Dinner?", *Time*, 161(20), 85.
Jöreskog, K G., Sörbom, D. Du Toit, S. and Du Toit, M. (2000), *LISREL 8: New Statistical Features*, (2nd Edition), Chicago, IL: Scientific Software International, Inc.
Kaynama, S.A. and Black, C.L., (2000), "A Proposal to Assess the Service Quality of Online Travel Agencies: An Exploratory Study", *Journal of Professional Services Marketing*, 21(1), 63-72.
Kelley, S. W. and Davis, M.A., (1994), "Antecedents to Customer Expectations for Service Recovery," *Journal of the Academy of Marketing Science*, 22 (1), 52-61.
Kettinger, W.J., Lee, C.C. and Lee, S. (1995), "Global Measures of Information Service Quality: A Cross National Study", *Decision Sciences*, 26 (5), 569-588.
Klassen, R.D. and Jacobs, J., (2001), "Experimental Comparison of Web, Electronic and Mail Survey technologies in Operations Management", *Journal Of Operations Management***,** 19 (6), 713-728
Koch, J.V. and Cebula, R.J., (2002), "Price, Quality and Service on the Internet: Sense and Nonsense", *Contemporary Economic Policy*, 20 (1), 25-37.
Laseter, T., Berg, B. and Turner, M., (2003) "What FreshDirect Learned from Dell", *Strategy & Business*, Spring, 1-6.
Lovelock, C.H., (1983), "Classifying Services to Gain Strategic Marketing Insights," *Journal of Marketing*, 47 (Summer), 9-20.
Malone, T.W., Yates, J., and Benjamin, R.I., (1987), "Electronic Market and Electronic Hierarchies", *Communications of the ACM*, 41(1), 73-80.
McDonald, R.P. and Marsh, H.W. (1990), "Choosing a Multivariate Model: Noncentrality and Goodness of Fit," *Psychological Bulletin*, 107 (2), 247-255.
McLaughlin, K., (2003), "Back from the Dead: Buying Groceries Online", *Wall Street Journal*, Feb 25,D.1.
Meuter, M.L., Ostrom, A.L., Roundtree, R.I. and Bitner, M.J., (2000), "Self-service Technologies: Understanding Customer Satisfaction with Technology-based Service Encounters", *Journal of Marketing*, 64, (3), 50-64.
Mnyandu, E., (2003), "Online Grocers Mum on Profits", *Chicago Tribune*, February 8.
Papke-Shields, Karen E., Manoj K. Malhotra and Varun Grover, "Strategic Manufacturing Planning Systems and Their Linkage to Planning System Success", *Decision Sciences*, 33 (1), 1-30.
Parasuraman, A., Zeithaml, V.A. and Berry, L.L., (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49 (Fall), 41-50.
Podsakoff, P.M. and Organ, D.W. (1986), "Self-Reports in Organizational Research: Problems and Prospects," *Journal of Management*, 12 (4), 531-544.
Rizzo, T., (2001), "The Death of Webvan", *Internet World*, August 1, 4-5.

Shostack, G.L., (1977), "Breaking Free from Product Marketing," *Journal of Marketing*, 41 (April), 73-80.
Skinner, W., (1969), "Manufacturing: missing link in corporate strategy", *Harvard Business Review*, 47(3), 136-145.
Stassen, R.E. and Waller, M.A., (2002), "Logistics and Assortment Depth in the Retail Supply Chain: Evidence from Grocery Categories", *Journal of Business Logistics*, 23(1), 125-143.
Tanskanen, K., Yrjola, H. and Holmstrom, J., (2002), "The Way to Profitable Internet Grocery Retailing – Siz Lessons Learned", 30 (4), *International Journal of Retail & Distribution Management*, 169-178.
Teas, R. K., (1993), "Expectations, Performance Evaluation, and Consumers' Perception of Quality," *Journal of Marketing*, 57 (October), 18-34.
Van Dyke, T. P., Prybutok, V.R. and Kappelman, L.A., (1999), "Cautions On the Use of the SERVQUAL Measure to Assess the Quality of Information Systems Services," *Decision Sciences*, 30 (3), 877-891.
Vickery, L. and Agins, T., (2001), **"**Retailers Find Web Apparel Unprofitable", *Wall Street Journal*, June 25, B.6.
Wingfield, N., (2003), "Online Grocery Shopping is picking up Speed as Shoppers' Fears Ease", *St. Louis Today*, February 28.
Yrjola, H., (2001), "Physical Distribution Considerations for Electronic Grocery Shopping", *International Journal of Physical Distribution & Logistics Management*, 31(10), 746-761.
Zeithaml, V.A., Berry, L.L. and Parasuraman, A. (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (April), 31-46.