

INTERNET VS. TELEPHONE DATA COLLECTION: DOES METHOD MATTER?

Research Findings

Differences in survey responses resulting from using an Internet survey instead of a phone survey were explored in a recent Burke, Inc. research and development project.

Three-hundred eighty-six Web interviews and 500 telephone interviews were conducted for the study. Unlike previous efforts to validate the Web as an effective medium for conducting surveys, this project attempted to isolate differences in responses directly related to the survey method. People taking both the Web and telephone surveys were recruited over the phone using a random-digit dialing technique. Only those with access to the Internet were given an opinion to participate in the study, and individuals were randomly assigned to take the survey using the Web or over the telephone.

This white paper presents findings from the research. Topics include:

- The inefficiency of phone calling as a recruitment vehicle for Web surveys;
- Differences in respondent characteristics that emerge despite the random assignment of individuals to take Web or phone surveys;
- Respondent views about the survey process itself and the effect of including a “don’t know” option for some Web respondents;
- The use of the Web compared to the phone for gauging opinions about sensitive issues;
- Scale usage by Web respondents and phone respondents, especially as it relates to concept testing;
- The use of the Web compared to the phone for conjoint studies; and
- The use of the Web compared to the phone for customer satisfaction studies.

PHONE CALLING INEFFICIENT IN RECRUITING WEB SURVEY RESPONDENTS

Using random digit dialing does not appear to be an effective method of recruiting

people to take Internet surveys — even when offering a \$10 incentive.

Using phone recruitment for a Web interview yielded an extremely inefficient net completion rate of just 20%:

- Only 77% of respondents with Internet access randomly selected to take a questionnaire over the Web agreed to do so during the phone recruitment;
- Only 55% of those who agreed to take the questionnaire actually went to the Web site;
- Only 85% of those people who went to the Web site completed any of the interview; and
- Only 56% of those who started the interview actually completed it.

By contrast, 86% of those with Internet access randomly selected to take the phone interview did complete it. This is despite the fact that the study took longer to complete over the phone. The average completion time on the Web was 12.5 minutes, compared to 19.4 minutes over the phone.

There is also lag time between phone recruiting and receiving completed Web interviews. It took, on average, 3.1 days from the time that Web respondents were recruited for them to complete the survey. For respondents who opted to write down the survey’s Web address during the phone



recruitment, the average time to get a completed survey was 1.8 days.

For respondents who opted to receive an emailed invitation with a hyperlink, the average time to get a completed survey was 3.6 days. This, however, includes a short delay between recruitment and Burke sending the email to potential respondents.

Another factor making phone recruitment unattractive is that it is still only a minority of American households that actually have access to and use the Internet. Burke’s findings relating to Internet access are similar to those of other surveys taken in the Spring of 1999. Burke found that, among all households with a male or female at least 18 years of age, 38% indicated that they have at least one member with access to the Internet, either at home or elsewhere. Of those households with access to the Internet, 91% indicated that they have at least one member who actually uses the Internet. Therefore, among all households with a male or female at least 18 years of age, 35% indicated that they have at least one household member who has access to and uses the Internet.

Burke discovered, however, that only 65% of those asked to participate in the Internet survey said they had email. Because Internet panels typically use email to request participation in studies, it appears as if their access is limited to approximately 23% of the total population (35% with Internet access X 65% with email). Also, panel respondent cooperation rates are typically around 50%; this self-selection makes it even more difficult to conclude that Internet panel respondents are representative of the entire population.

WEB SURVEY RESPONDENTS MORE EXPERIENCED INTERNET USERS THAN PHONE RESPONDENTS

Even when survey respondents who use the Internet are randomly assigned to phone interview and Web interview groups,

demographic and other differences still emerged among those who actually completed the interviews.

The demographic differences found between respondents completing the two survey versions included differences in formal education levels and area of residence. Web respondents were much more likely to be college grads (57% vs. 46%), while phone respondents were much more likely to have not attended college (26% vs. 13%). Web respondents were much more likely to be living in a suburban area (51% vs. 42%), while phone respondents were much more likely to be living in rural areas (35% vs. 25%).

Perhaps the key difference, however, is that respondents who will actually complete a Web interview tend to be more experienced Web users. On average, Web respondents spend more time online per week than those in the phone group (10.6 hours per week vs. 7.9 hours per week). They also report having been online for a somewhat longer period of time (a mean of 19.34 months vs. 18.55 months) See Figure 1, top right.

In addition, a greater proportion of Web respondents reported accessing the Internet from locations other than the one they use most often (about 65% vs. 49% of phone respondents). These locations included friends' and relatives' homes, as well as airports and hotels.

Respondents were also asked whether or not they had participated in 15 different online activities. Web respondents reported performing, on average, about 11 actions online, compared with only 8 actions among phone respondents. In particular, Web respondents are more likely than phone respondents to have purchased something using the Internet in the past six months. Respondents in the Web group are significantly more likely than those in the phone group to say that they had purchased books, CDs or tapes, airline tickets, hotel reservations, stocks/bonds, and rental cars. See Figure 2, bottom right.

Some differences also appeared when respondents were asked to respond to "early adopter" attitude scales. These were seven-point scales with verbal endpoints, with a "1" indicating that the respondent disagreed completely with the statement, and "7" indicating that the respondent agreed completely.

Web respondents were more likely than phone respondents to express agreement with the statements, "I am used as a source

Figure 1: Amount of Time Since First Being Online (% of Respondents)

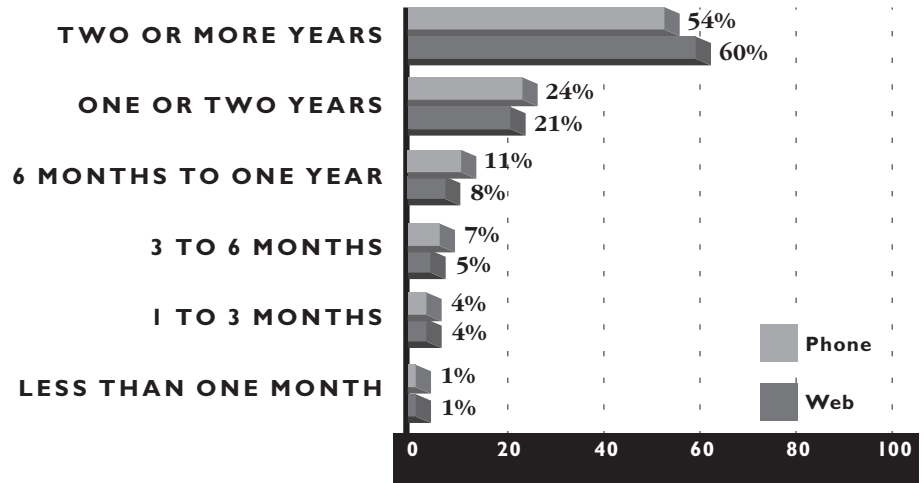
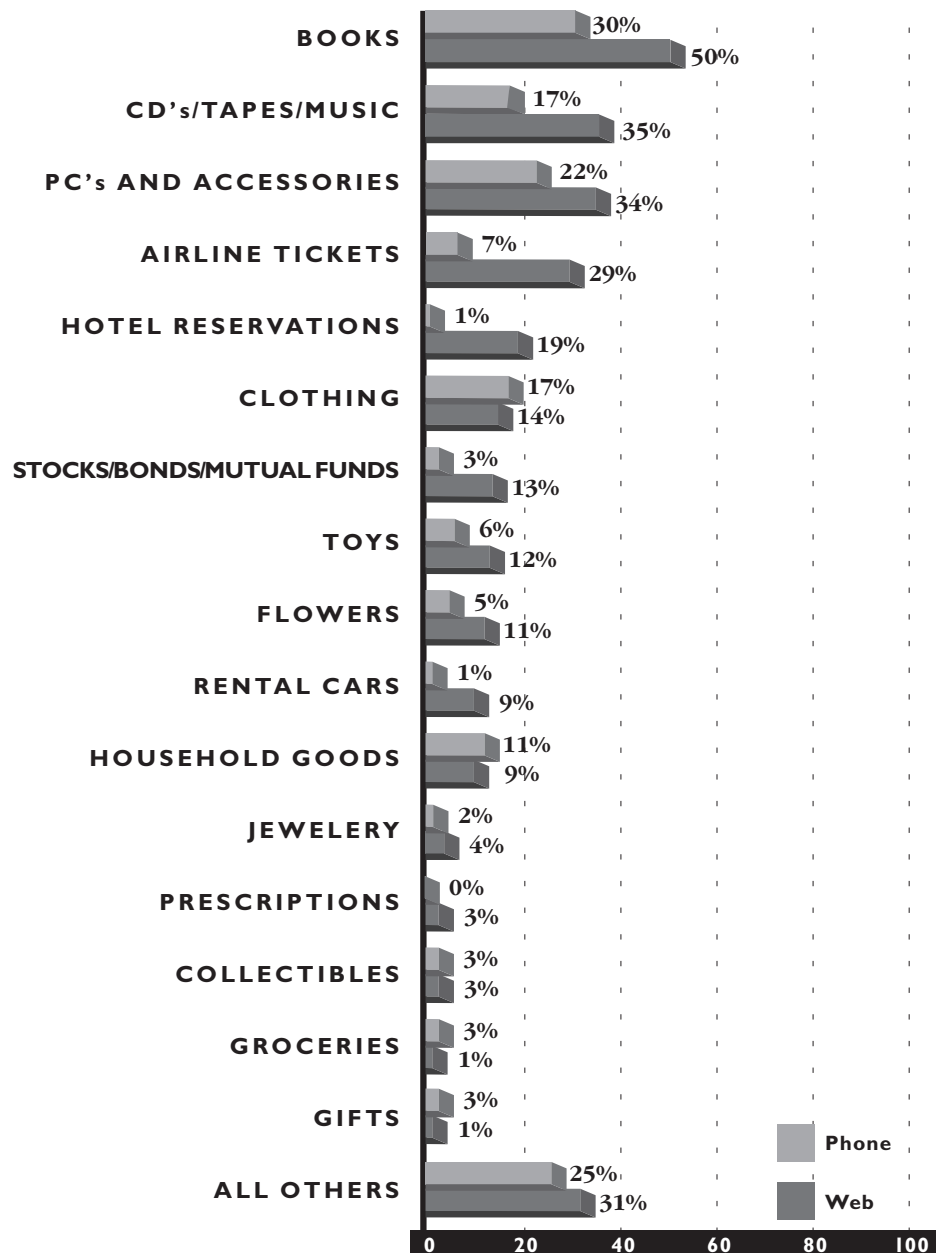


Figure 2: Purchasing on the Web (% of Respondents)



of advice for computer-related products by friends,” and “For computer-related products, my friends are likely to ask my opinions.” Web respondents were less likely than phone respondents to express agreement with the statements, “Buying a new product that has not yet been proven is usually a waste of time and money,” and “If people would quit wasting their time experimenting, we would get a lot more accomplished.”

In a sense, by asking respondents to complete a survey over the Web, Burke was “calling the bluff” of those who indicated that they had used the Internet. Presumably, those who indicated that they had used the Internet, but in reality had very little experience with it were less likely to complete the survey than skilled Web users. Phone respondents, by contrast, did not have to prove their ability to access and perform tasks on the Web. This could explain the discrepancies found.

WEB RESPONDENTS EXPRESS WILLINGNESS TO PARTICIPATE IN FUTURE SURVEYS
Including “Don’t Know” Option Does Not Increase Survey Completion Rate or Respondent Enjoyment

Web respondents indicate more willingness than phone respondents to participate in future marketing research surveys.

About one-third (33%) of those taking the survey on the Internet indicated that they “definitely would” participate in future surveys. By contrast, only 18% of those taking the survey over the phone indicated that they “definitely would” participate in such a study again.

Web respondents who were not given a “don’t know” option in their surveys were especially likely to react positively to taking the survey. Almost nine out of ten (89%) of Web respondents not given a don’t know option called the survey “very” or “somewhat” enjoyable, compared to just 70% of Web respondents given a don’t know option. Similarly, 88% of Web respondents not given a don’t know option called the survey “very” or “somewhat” interesting, compared to just 75% of Web respondents given a don’t know option.

Respondents asked to take the survey over the Web were randomly assigned to two different surveys, one with don’t know options in response sets and the other without. Two-hundred twelve responses were completed by respondents assigned to take a survey with the don’t know option, and 174 responses were completed over the

Web by people assigned to take the survey without the don’t know option.

Overall, survey findings suggested that, in most cases, there is not a compelling reason to include don’t know options in Web surveys. If provided such an “out,” many respondents will take it. “No response” rates for the study as a whole were significantly higher among Web respondents given this option on the screen (3.0%) than among phone respondents (1.4%). This was especially true on questions relating to social/political issues (15.3% of Web respondents with a don’t know option vs. 6.8% of phone respondents). Differences in no response rates may stem from the fact that phone respondents are allowed to give a “don’t know” answer, but are not prompted for one. A portion of the Web respondents, on the other hand, were given a “don’t know” choice along with their other response choices.

If not given this “out,” respondents do tend to answer the question. There was no greater tendency for Web respondents to terminate the interview if they were not provided “don’t know” options.

The difference in phone and Web respondents’ willingness to participate in future surveys perhaps is a result of phone respondents finding the interview somewhat intrusive, while Web respondents see themselves as having more control over the interview. In addition, the longer average length of the phone interview (19.4 minutes) compared with the Web interview (12.5 minutes) might have contributed to the observed difference. It is unclear, however, why Web respondents not given the “don’t know” option reacted more positively to the survey than Web respondents who were given this choice.

WEB SURVEYING CAN BE AN EFFECTIVE METHOD FOR GAUGING OPINIONS ON SENSITIVE ISSUES

The Web can be an excellent avenue for measuring opinions on sensitive issues. If, however, there is a refusal option provided, “no response” rates can run very high on the Web. As a result, study findings can be packaged very differently depending upon whether or not a “don’t know” option is included.

Six social/political questions were included on the Burke survey to test reactions to sensitive issues. These included questions on respondents’ positions on the legality of

abortion and the death penalty; identification as “liberal,” “conservative,” or “somewhere in the middle”; political party affiliation; belief in God; and belief in life after death. On all six items, the “no response” rate was higher among Web respondents given the option than it was among phone respondents. On five of the six items the difference was statistically significant at the 95% confidence level.

The abortion issue provides the best example of how the presence of a “no response” alternative on the survey could change how findings are packaged. Among Web respondents who were given such an option, 23% chose to take it. Forty-five percent of the group of Web respondents given the “no response” option indicated that they were in favor of abortion being legal, while 32% indicated that they were opposed. Among Web respondents who were not given a “no response” option, 64% indicated being in favor of abortion being legal, while 36% indicated being opposed. Whether or not a majority of respondents favors abortion being legal therefore depends upon whether or not a “don’t know” option is included. (Among phone respondents, 54% indicated that they favored abortion being legal, while 35% indicated opposition, and 11% gave no response.)

Similarly, the percentage of Web respondents indicating that they were opposed to the death penalty differed significantly depending on whether or not a “no response” option was given. When no such alternative was presented, 28% indicated that they oppose the death penalty, with 72% indicating support. With such an option, only 15% indicated opposition, with 64% indicating support. Among phone respondents, 23% indicated opposition, 63% indicated support, and 15% gave no response.

When asked to describe themselves “liberal,” “conservative,” or “in the middle,” the percentage describing themselves as “in the middle” was significantly higher among phone respondents (47%) than it was among Web respondents as a whole (39%). This suggests that people in a phone interview situation, in which they are speaking to another human being, might be less inclined than Web respondents to express an extreme position and risk “offending” the interviewer. Thirty-five percent of Web respondents described themselves as “conservative,” with 23% describing themselves as “liberal.”

Among phone respondents, 32% described themselves as “conservative,” with 19% describing themselves as “liberal.”

With the exception of the “no response” option, there were not significant differences between phone respondents and Web respondents as a whole in political party affiliation. Among phone respondents, 35% identified themselves as Republicans, 32% as Democrats, 22% as Independents, and 7% as members of some other party. Among Web respondents, 37% identified themselves as Republicans, 31% as Democrats, 20% as Independents, and 5% as members of some other party. In recent national polls of all Americans (not just those who have used the Internet), more people have described themselves as being Independents than as being Democrats or Republicans and no more than 2% have indicated belonging to some other party.

Belief in God and belief in life after death were similar among phone respondents and Web respondents when the responses of those having the “don’t know” option and those not having the “don’t know” option were combined. Among phone respondents, 90% indicated they believe in God and 77% indicated they believe in life after death. Among all Web respondents, 87% indicated they believe in God and, again, 77% indicated they believe in life after death.

SCALE USAGE CAN VARY DEPENDING UPON WHETHER TELEPHONE OR WEB SURVEYING IS USED

Concept-Testing Studies Among Research That Can Be Affected

Depending on the type of scale used, responses to Web survey items can vary considerably from responses to phone survey items. In particular, when scales for survey items are completely verbally labeled, phone respondents are significantly more likely than Web respondents to use scale endpoints.

The differences between scale usage in Web and phone responses can be most clearly seen when examining responses to purchase-intent items. There were four items on the Burke study relating to how likely respondents would be to purchase a described new product concept, with the possible responses being that they “definitely would,” “probably would,” “might or might not,”

“probably would not,” or “definitely would not.”

Among phone respondents, 45%, on average, indicated that they “definitely would” or “definitely would not” purchase the concepts described. Among Web respondents, only 27%, on average, used one of the scale endpoints.

The tendency for more Web respondents to respond in the middle of the scale was true despite very different levels of popularity among the concepts described. Regarding a daily shower cleaner, 19% of phone respondents indicated that they “definitely would” purchase it and 23% indicated that they “definitely would not.” Among Web respondents, however, just 12% indicated that they “definitely would” purchase it and only

13% indicated that the “definitely would not.” See Figure 3 below.

Regarding a DIVX player (a system that would allow consumers to play “expiring” digital video discs instead of renting videos), 2% of phone respondents indicated that they “definitely would” purchase it and 42% indicated that they “definitely would not.” Among Web respondents, only 1% indicated that they “definitely would” purchase the product and just 26% indicated that they “definitely would not.” See Figure 4 below.

It is noteworthy that, in each of these cases, if the scales were converted to a 1-5 numerical scale, the “mean” responses would not substantially differ between methods. (For the daily shower

Figure 3: Likelihood to Purchase a Daily Shower Cleaner (% of Respondents)

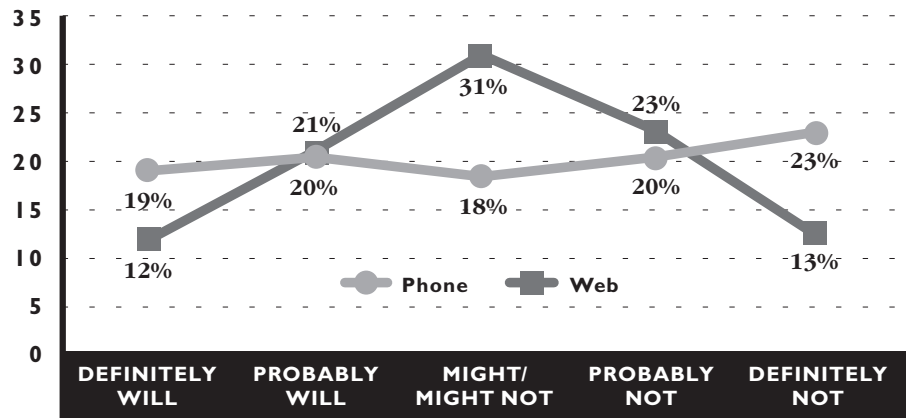
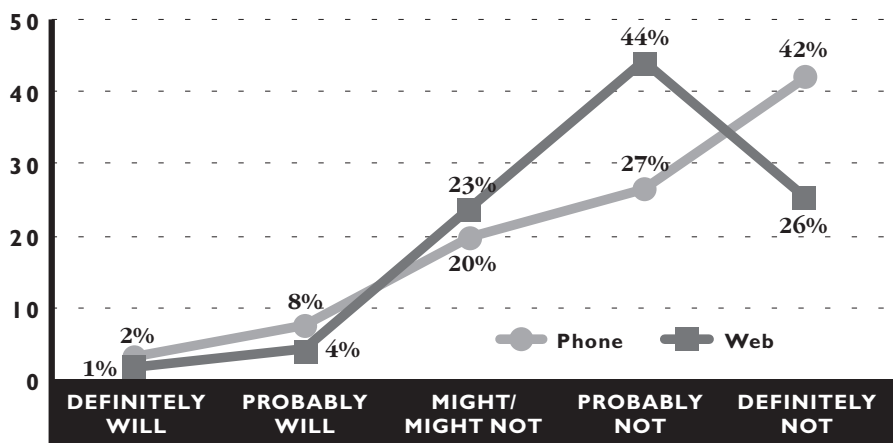


Figure 4: Likelihood to Purchase a DIVX Player (% of Respondents)



cleaner, the mean response would be 2.92 among phone respondents vs. 2.96 among Web respondents; for the DIVX player, the mean response for phone respondents would be 2.01 and for Web respondents would be 2.09.) Typically, however, responses to purchase intent scales are adjusted downward in order to create demand estimates. The differing use of scales by phone and Web respondents suggests that some adjustment schemes would result in very different forecasts depending upon whether the survey took place over the Web or over the phone.

The study findings do not indicate why use of endpoints would be greater in phone surveys than in Web surveys. Responses could be affected by the fact that in phone surveys, possible responses are read by an interviewer, whereas in Web surveys, respondents can see all possible responses at the same time. Respondents hearing responses read to them might focus more on the first and last ones named, while Web respondents might be visually drawn to responses closer to the middle of the scale.

Other survey items where this pattern was seen included items relating to price sensitivity. When asked how likely they would be to continue shopping at their grocery store if prices increased 10%, 3% of phone respondents indicated that they “definitely would” and 41% indicated that they “definitely would not.” Among Web respondents, 2% indicated that they “definitely would” and 34% indicated that they “definitely would not.”

Scale usage did not substantially differ between phone respondents and Web respondents when, instead of a completely verbally labeled scale, a numerical scale with verbal endpoints was used. Figure 5, top right, shows the response patterns to an item using a 1-7 scale, where “1” indicated complete disagreement and “7” indicated complete agreement.

It is likely that a segmentation method using responses to such an item would derive similar solutions for both data collection methods. In a survey program in which there is a need to combine responses obtained through Web and phone surveys, using numerical scales with verbally anchored endpoints might serve as the best approach.

PHONE CONJOINT RESULTS AND WEB CONJOINT RESULTS CAN BE NEARLY IDENTICAL

With the proper scale, conjoint data collection over the Web or over the phone can yield nearly identical results.

The conjoint exercise in the Burke study asked respondents to rate their likelihood to purchase side items from quick-service restaurants. Ratings were given on a numerical five-point scale, with “1” indicating “not at all likely to buy” and “5” indicating “very likely to buy.” Use of a numerical scale with verbally anchored endpoints was important, as other portions of the study indicated that Web and telephone responses can differ considerably when a completely verbally labeled scale is used.

Three types of side items (french fries, onion rings, hash browns), restaurant chains (McDonald’s, Wendy’s, Burger King), and price levels (\$1.29, \$1.09, and 89 cents) were asked about in the exercise. A balanced and orthogonal fractional design was employed in which each respondent was exposed to nine of the 27 possible configurations. Responses were also collected for a 10th “hold-out” profile, which was not included in utility estimations.

Patterns of responses to the profiles seen were very similar for Web and phone respondents. For all phone respondents, 24% of the profiles were rated a “5,” 16% a “4,” 21% a “3,” 14% a “2,” and 25% a “1.” For all Web respondents, 24% of the profiles were rated a “5,” 15% a “4,” 20% a “3,” 15% a “2,” and 26% a “1.” See figure 6 below.

Figure 5: My Friends Think of Me as a Good Source of Information When It Comes to New Products or Sales (% of Respondents)

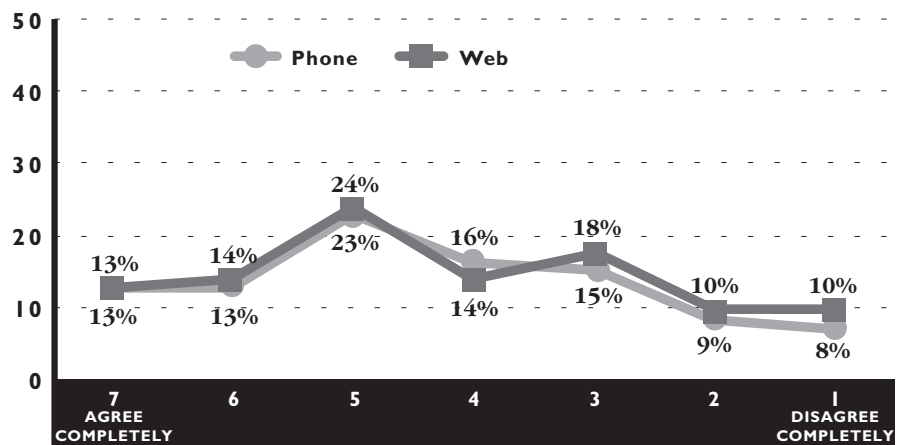
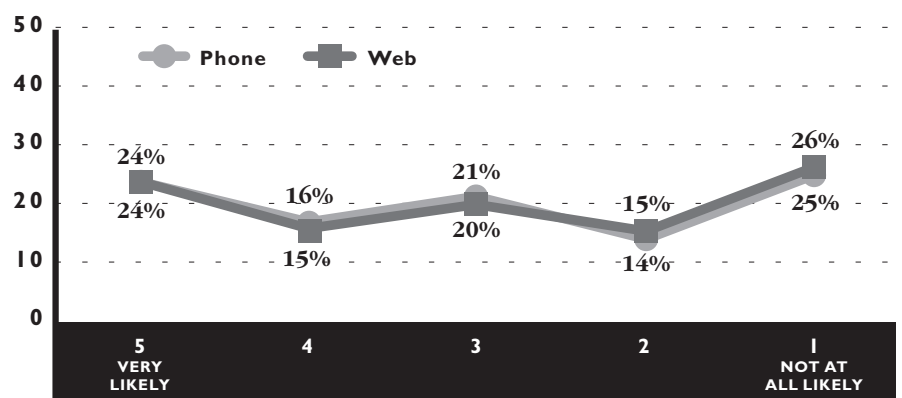


Figure 6: Conjoint Exercise Scale Usage (% of Respondents)



Lower price levels were almost equally favored by Web and phone respondents. Similarly, both Web and phone respondents indicated a preference for McDonald's over Burger King and Wendy's. Although french fries were the preferred side item by both Web and phone respondents, Web respondents did express a preference for onion rings over hash browns, while phone respondents expressed a slight preference for hash browns over onion rings. See Figure 7, below.

The range of utility values across each category — often used as a measure of the category's overall importance — was very similar for Web respondents and phone respondents. R-square values, indicating the consistency across the nine responses

given by each respondent, were also similar (0.54 for phone respondents, 0.51 for Web respondents). The correlation coefficient between actual ratings of the "hold-out" profile and ratings "predicted" by responses to the other nine items was 0.59 for phone respondents and 0.62 for Web respondents.

One data problem that can arise in conjoint analyses is having a portion of respondents who give all profiles the same rating. This happened among 5% of the phone respondents. For respondents performing the conjoint exercise on the Web, the percentage was only slightly higher (6%).

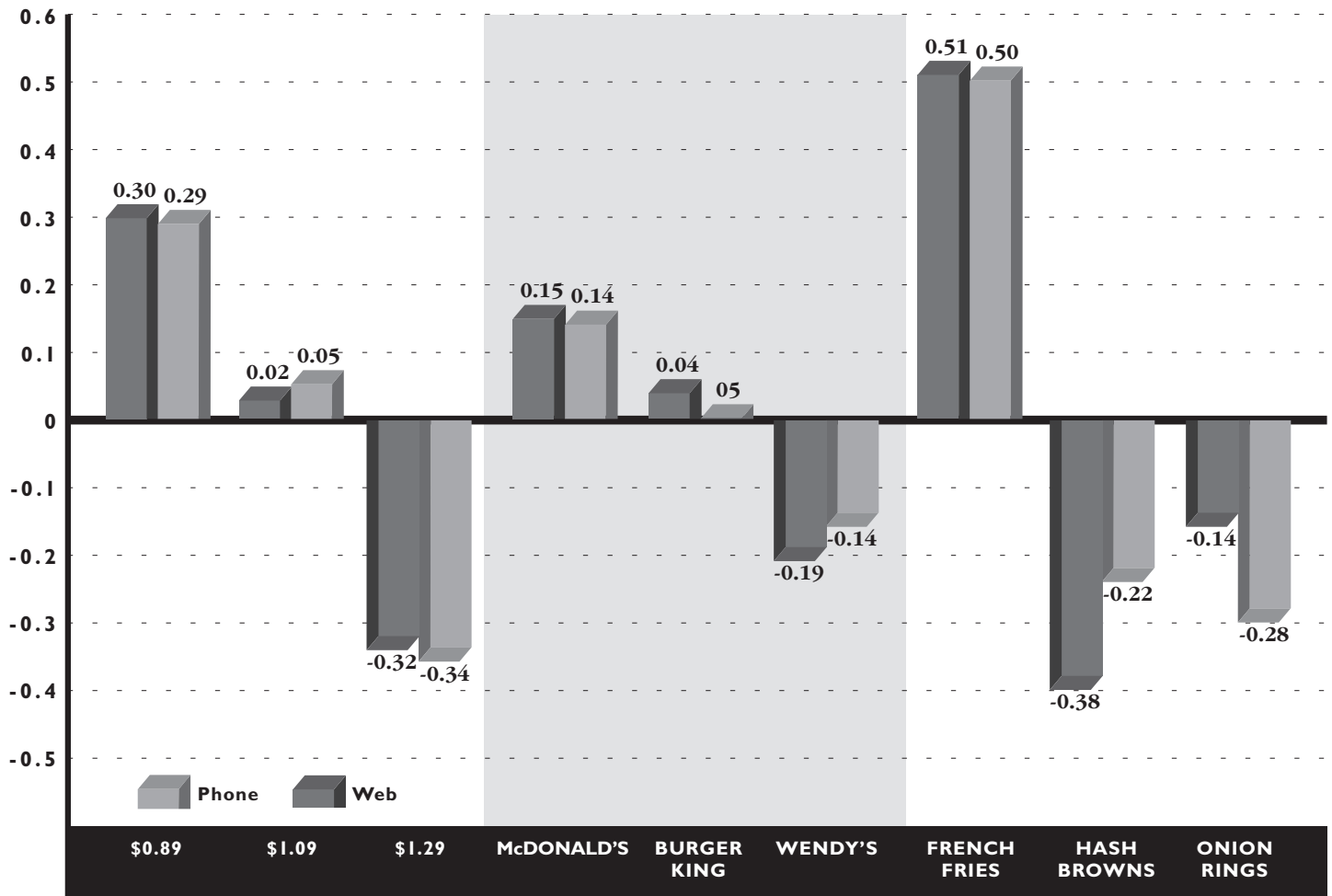
Burke's findings on this exercise suggest that researchers should feel free to use

the Web as a method for conducting conjoint studies, as long as those who use the Internet are considered representative of the survey population. The study also suggests that it should be possible to mix phone and Web responses to conjoint exercises as long as numerical scales with verbally anchored endpoints are used for ratings.

WEB STUDIES CAN PRODUCE THE SAME CUSTOMER SATISFACTION FINDINGS AS PHONE STUDIES

Identical measurements of overall customer satisfaction and loyalty can be obtained through Web and phone surveying. Consistent improvement recommendations can also be gleaned from data obtained over the phone and over the Web.

Figure 7: Conjoint Exercise Derived Utility Values



The customer satisfaction portion of the research project, which asked respondents to rate the grocery store at which they shop most often, used Burke Customer Satisfaction Associates' Secure Customer Index® (SCI®) as a measure of overall satisfaction and loyalty. The SCI indicates the percent of "Secure Customers," defined as those who are "very satisfied" with a store, "definitely would" recommend it to others, and "definitely will" shop there again.

The study found that 43% of both Web respondents and phone respondents could be classified as Secure Customers of the grocery store they shop at most often. The identical results suggest that customer satisfaction tracking studies using SCI as an overall measure should be able to move from phone-based data collection to Web-based data collection without having a substantial impact on scores. This finding also opens up the possibility of performing studies that use both the phone and the Web as data collection techniques and combining findings from the two approaches.

Some caution, however, should be exercised before taking such steps. Although SCI scores in this study were identical using the two methods, in other respects customer satisfaction results can be impacted by a difference in survey method. Possible impacts include:

- *Differences in classification of customers who are not identified as "Secure."* The Burke CSA approach classifies respondents who are not deemed "Secure" as being either "Favorable," "Indifferent," or "At Risk." Differing scale usage by Web and phone respondents might result in fewer At Risk respondents and more Indifferent respondents when surveys are done using the Web.
- *Differences in derived importance measures of attributes.* If the importance of attribute performance is measured by its correlation with customer security, these scores may tend to be higher if respondents use the Web instead of using the phone. This is particularly true if respondents using the Web are not given a "don't know" option.

The differing percentages of respondents categorized as Indifferent and At Risk

results from differing uses of completely verbally labeled scales by Web survey respondents and phone survey respondents. Burke's research indicates that phone respondents have a greater tendency in these cases to use scale endpoints than do Web respondents.

The three component questions used to calculate the SCI are completely verbally labeled scales. Respondents are asked first if they are "very satisfied," "somewhat satisfied," "somewhat dissatisfied," or "very dissatisfied" with a store. They are then given five possible responses relating to both recommending the store and shopping there again: "definitely," "probably," "might or might not," "probably not," and "definitely not."

For the grocery store study, phone respondents used SCI component scale endpoints 63% of the time, while Web respondents used SCI component scale endpoints just 56% of the time. Because respondents are classified as At Risk if they give the lowest score on any one of the three scales, a greater percentage of phone respondents were classified as At Risk.

Greater use of scale endpoints did not lead to more phone respondents being classified as Secure because Web respondents displayed somewhat more consistency in their scale use than did phone respondents. Cronbach's alpha, a statistical measure of internal consistency, was .75 for phone respondents, .85 for all Web respondents, and .89 for Web respondents assigned to take a version of the survey that did not have "don't know" as a response option.

The higher internal consistency among Web responses also leads to the higher derived attribute importance scores associated with Web surveys.

The customer satisfaction portion of the research project asked respondents to rate the importance of 13 attributes relating to grocery stores, as well as their grocery store's performance on these attributes.

Attribute importance and performance were both rated on a 1-9 scale with verbally anchored endpoints.

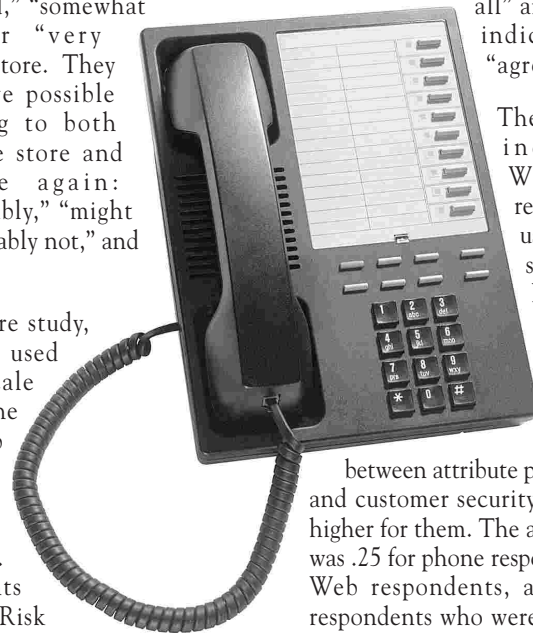
Respondents were first asked to what extent they agreed that any grocery store must be, for example, clean and well-lit. They were then asked to what extent they agreed that the grocery store they shopped at most often was clean and well-lit. A "1" response indicated that they "did not agree at all" and a "9" response indicated that they "agreed completely."

The research project indicated that Web and phone respondents tend to use such scales very similarly. Because, however, there more internal consistency displayed by Web respondents, the correlation

between attribute performance ratings and customer security was, on average, higher for them. The average correlation was .25 for phone respondents, .31 for all Web respondents, and .34 for Web respondents who were not given "don't know" options.

It should be noted that extremely high internal consistency on survey responses is not necessarily desired. Multiple questions are asked in surveys because it is assumed that responses will differ from question to question.

When respondents are not allowed to answer a question by indicating that they "don't know," they might be forced to rely upon their general impressions of, for example, the grocery store when responding to a survey. One of the attributes included in this study was the availability of in-store banking at the grocery store. Respondents who truly did not know about the availability of in-store banking were nevertheless forced to offer some response in order to complete the survey. Perhaps the correlation between attribute performance ratings and customer security was higher because they drew upon their overall feelings about the store when crafting a response.



Impact scores were, however, fairly consistently higher across the set of attributes, so issue prioritization findings did not change substantially across methods.

Figure 8, top right, describes the prioritizing of attributes based on a grid where the vertical axis represents the impact scores and the horizontal axis represents attribute performance relative to requirements.

Figure 9, bottom right, identifies the classification of each attribute using this prioritization scheme.

Issues are identified as "Critical Issues" that should be priorities for performance improvement when the performance ratings relative to customer requirements are below average and the correlation between performance ratings and customer security is above average. Among both Web and phone respondents, the same issues were identified as priority issues for improvement: having enough checkout lanes, having fresh produce daily, having a good selection of fresh meats and fish, and having a wide variety of products and brands.

The similarity in recommendations arising from the differing methods again suggests that, although some caution should be used, phone-based customer satisfaction studies could be transferred in whole or in part to the Web without a loss of continuity.

The survey was conducted from April 5, 1999, to May 9, 1999. Qualified respondents were at least 18 years old and had access to and used the Internet. An incentive of \$10 was offered to all participants.

For more information, please contact:

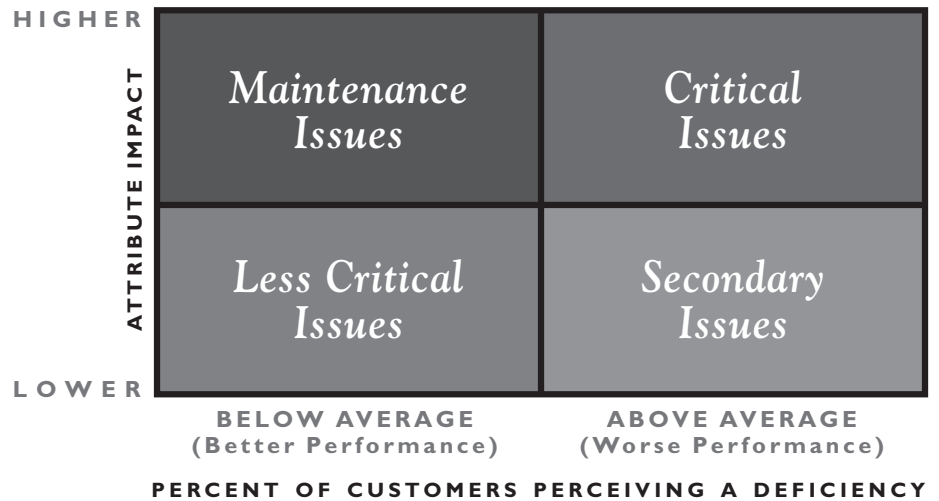
Jeff Miller
Senior Vice President, Burke Interactive
513/684-7626
jeff.miller@burke.com

or

Allen Hogg
Director of Marketing, Burke Interactive
203/431-5906
allen.hogg@burke.com

© 1999 Burke, Inc. All rights reserved.

Figure 8: Issue Prioritization Grid



**Figure 9: Classification of Attributes
Phone Respondents Vs. Web Respondents**

	PHONE	WEB
ENOUGH CHECKOUT LANES	Critical	Critical
FRESH PRODUCE DAILY	Critical	Critical
GOOD SELECTION OF FRESH MEATS AND FISH	Critical	Critical
WIDE VARIETY OF PRODUCTS AND BRANDS	Critical	Critical
CLEAN AND WELL-LIT	Maintenance	Maintenance
DELI WITH BROAD SELECTION	Maintenance	Maintenance
WIDE VARIETY OF ETHNIC FOODS	Secondary	Maintenance
AISLES WIDE ENOUGH	Less Critical	Secondary
WHEEL CHAIR SHOPPING CARTS AVAILABLE	Less Critical	Less Critical
CONVENIENT LOCATION	Less Critical	Less Critical
FLORIST WITH FRESH FLOWERS	Less Critical	Less Critical
IN-STORE BANKING FACILITY	Less Critical	Less Critical
STORE-BRANDED PRODUCTS	Less Critical	Less Critical



BURKE INCORPORATED
805 CENTRAL AVENUE · CINCINNATI, OH 45202 · 1-800-264-9970
www.burke.com