# THE ROLE OF THE INTERNET IN SURVEY RESEARCH: GUIDELINES FOR RESEARCHERS AND EXPERTS

**Thomas J. Maronick**
**Towson University, Towson, USA**

## ABSTRACT

*The paper examines the relationship between web-based surveys and more traditional survey methods. The author examines issues related to: sample quality, response rates, sample panels, and the use of incentives in a web-based environment. The author also discusses web-based survey design issues and provides guidelines for effective utilization of web-based surveys in litigation and research contexts. The author identifies effective implementation techniques, including use of devices to assure the quality of the resulting data.*

## INTRODUCTION

The role of the Internet as a vehicle for communications is no longer questioned. Within the US, estimates of Internet usage range from 69% (Pew Research 2007) to 75% (Fadner and Mandise 2004) and over 40% worldwide, with growth in European and Asian markets growing at double-digit rates (Roster 2004). Moreover, while studies have shown lower coverage levels for Hispanics and residents of rural areas, these differences are diminishing with every passing year (Zang 2000). Additionally, as noted by Cambiar (2006) "online research continues its headlong march to become one of the most dominant (if not the most dominant) data collection methodology worldwide." Evidence that the Internet has become the vehicle of choice for many survey research organizations is found throughout the literature. For example, Internet research now represents 10-30% of customized research for AC Neilson and almost 47% of Harris Interactive's polling research (Einhart 2003). In fact, estimates are that 87% of research companies use online methodologies (Cambier 2006), with an estimate that over 35% of data collected in 2006 was on-line. (Bradley 2006).

Importantly, web-based survey methodologies have begun to move outside the corporate world into the domain of the court system, with judges across the U.S. being asked to consider web-based surveys as part of deceptive advertising and trademark litigation. As a consequence, lawyers and experts are being asked to defend and justify web-based research much like when telephone surveys were introduced in the 1960's and mall-based studies entered the courtroom in the 1980's.

Research evaluating on-line survey methodologies has also begun to emerge, with studies comparing differences in response rates and data quality between mail and/or phone surveys and Internet or web-based methods. Studies have also examined the impact of incentives, types of contacts or invitations used in web-based studies, and question format on data quality and willingness to respond.

The question is, what guidelines should a researcher or expert consider when deciding to utilize a web-based survey. This paper sets out those factors which are most likely to be raised when evaluating online research, either as part of an organization's corporate strategy or by judges and attorneys being asked to consider online research results as part of a litigation strategy.

phone or mail studies and, generally, are lower cost, particularly if the research firm has a targeted sample with email addresses (Bransberger et al 2007).

Internet surveys are not without disadvantages. Foremost is the question of the representativeness of the sample and, therefore, the projectability of the results to the population, even of a targeted population, as, for example, in Business to Business studies. However, even it a B2B environment, while firms may have email addresses of current customers, many have had difficulty obtaining representative samples of potential respondent/customer email addresses. Moreover, questions have been raised about whether the quality of data derived in an Internet survey is sacrificed because of the speed with which the data are collected. Finally, while the total number of responses can be increased simply by increasing the Internet survey sample, questions have been raised about actual response rates for Internet surveys (Couper 2000).

## Internet Availability

In 2007 approximately 71% of the U.S. population used the Internet, with 91% sending or receiving e-mail (Pew Research 2007). Importantly, a growing percent of those with Internet service have high-speed broadband technology, i.e., DSL or cable, making it possible for faster downloading of file containing audio and/or video images. The data also show there is no difference between men and women on Internet usage (Couper 2000). However, significantly fewer individuals over 65 use the Internet (32%) compared to those 18-49 (85%) and those 50-64 (65%). Finally, while there is little difference between White and Hispanic consumers on Internet use (73-78%), Blacks use of the Internet is somewhat lower (62%) (Pew Research 2007). Finally, research in 2006 has shown that two-thirds (68%) of those with incomes over $50,000 have broadband connections in their homes (Pew Research 2007). On the other hand, in the case of many lower income consumers, access to a computer may be limited to access through their public library.

For the marketing researcher considering an Internet survey, these data suggest the need to be assured that the target population for the survey research is one where a sufficient percentage had Internet access and, if the survey instrument includes graphics, that they have access to broadband technology for faster downloading.

## Interactive v Scrolling Surveys

There are two basic types of Internet surveys, interactive and scrolling. With interactive Internet surveys, respondents complete one screen at a time which is then transmitted to the server before the respondent "moves" to the next screen. This permits item-by-item real-time tracking of responses. With scrolling Internet surveys, respondents complete individual screens, then moves to the next screen (Dillman 2000, Couper et al 2001). The data are not transmitted to the server until the respondent has completed the entire survey (Peytchev et al 2006).

The benefit asserted for the interactive survey is that, once the respondent has finished a screen and the data transmitted, the respondent can't "go back," i.e., refer back to a prior question or response before answering a later question. It also gives the researcher greater control over skips and branching (Couper et al 2001). On the other hand, some have suggested that the interactive survey may take slightly longer to complete, depending on the download speed of their Internet connection, since the respondent has to wait for each response to download before proceeding. It has also been suggested that interactive surveys permit analysis of partial responses and permit respondents to leave the survey and come back at a later time and resume (Couper 2000). The criticism of the scrolling technology is that respondents can refer back to a prior question or response and that the researcher can't track partial responses.

The reality is that there really isn't any advantage, one way or the other for either methodology since, with the scrolling survey, the research designer has the option of limiting respondents' ability to refer back to a previous response when answering a subsequent question, even though the survey will not be transmitted until it is completed. Additionally, there really isn't much advantage to item-by-item real-time tracking over tracking completed questionnaires. Moreover, even with scrolling surveys, most Internet survey platforms permit analysis of partially completed surveys.

phone respondents (Couper 2000), analysis of "don't know/no opinion" responses for signs of "satisficing" "acquiescence" or "straight lining" responses (Fricker et al 2005), and utilization of different formatting techniques such as radio buttons as opposed to open-ended response options (Couper et al 2001). Other authors have examined data quality in terms of percent of negative and neutral responses. For example, Roster et al (2004), in a study comparing telephone and web responses, found that web surveys generated more negative and neutral responses and more extreme concerns than did their telephone survey. They also found there was no difference in the predictive validity between telephone and web-based surveys. Similarly, in a study of 3,238 respondents, the American Association for Public Opinion (AAPOR) and Harris Interactive found, "...high levels of reliability and validity that gives us much greater confidence that data collected from web-based surveys can be at least as good as data obtained in other research modalities, if not better" (AAPOR 2000).

## Sample Control

Another issue that has been raised related to Internet surveys is sample control, i.e., the ability to assure that the "right" individual responded and that multiple responses from the same respondent are eliminated (Couper 2000). While the issue of assuring the individual who actually completes the survey cannot be eliminated in Web survey any more than it could in mail studies, control questions may be possible to alert the researcher to questionable respondents. In the case of multiple responses from the same computer, this is rarely an issue with most Web surveys since most Internet survey platforms permit the researcher to restrict responses by limiting responses to one particular computer. This is done through the Machine Address Code, the unique identifier of each computer. Thus, once a response has been submitted from a particular computer, subsequent responses from that particular Machine Address Code or IP address are blocked (Frost and Sullivan 2007). On the other hand, there may be situations where multiple responses from the same computer is desirable as, for example, where a computer is set-up in a kiosk at a conference and attendees are asked to complete the on-line survey.

Another aspect of the sample control issue is the ability to track individual e-mail addresses of respondents. Most Internet survey platforms have e-mail address tracking as an option. However, while the option of tracking respondents may be desirable in mail surveys in order to reduce the cost of subsequent follow-up reminders and/or questionnaires, there is no real cost of following up in Internet surveys. It's simply a matter of sending out a second e-mail invitation. Thus, there really isn't a justification for tracking e-mail addresses of respondents, particularly when confidentiality is promised.

## Incentives

One significant difference between mail surveys and Internet surveys is the role of incentives. Numerous studies have shown that incentives, including cash and non-cash incentives, pre-payment and promises of payment, have resulted in higher response rates and higher quality results in mail surveys (Church 1993, Collins et al 2000). As noted above, many of the incentives that have traditionally been utilized to enhance response rates are not possible with Internet surveys. Thus, alternatives must be utilized. For example, Ray et al (2001), found in a study of Internet-based surveys that in 57% of the surveys respondents were promised survey results, while in 36% of the Internet surveys studied, the incentive that was a promised was inclusion in a draw/raffle. In another case, Cobanoglu and Cobanoglu (2003) found that offering a luggage tag to every respondent who completed an Internet survey resulted in a 31.4% response rate, compared to 23.9% for those in a control group who were not promised any incentive for participating.

## Panels

The final sample-related issue with Internet surveys is the use of pre-recruited panels as a sampling frame. Evidence suggests that firms are moving more toward the use of Internet panels for their surveys. For example, Cambiar (2006) found, in a survey of marketing research firms, that 66% were using outside panel companies for about a third (32%) of their research. Importantly, Black, CEO of Harris Interactive has said that their Harris Online Poll of 6.5 million members yields results that are no different from those derived from the traditional Harris telephone polls results. Moreover, since some of the Internet panels contain or the Internet survey firms have access to upwards of 2.5 million e-mail addresses (Zoomerang.com, Greenfield Online), their utilization permits pre-identification and selection of samples to fit desired demographics.

Similarly, virtually all Internet survey platforms include an option of rotating response categories in multiple response questions. This assures that order bias is eliminated in the study (Burns and Bush 2006). However, care must be taken when deciding on an Internet survey platform since some, but not all, permit holding one or more categories, such as "don't know" and "other" constant in the rotation or randomization.

## Random Assignment of Respondents to Treatments

One issue that frequently comes up is whether the Internet survey platform permits random assignment of respondents to treatments. This can be important in experimental studies, where one group of respondents is to see one exhibit (e.g., ad or package) and another isn't, i.e., a control condition. While it is an important technology, its availability isn't critical since alternatives are available. One is to ask respondents, at the critical branching point in the Internet survey, to indicate the last digit of their phone number (i.e., a random number). The research design then directs respondents to each branch based on their response.

## Branching and Piping

All telephone and face-to-face surveys include branching with "skip patterns." Similar technology is available with Internet surveys. While this permits more efficient survey designs, it may be a problem if multiple questions are asked together. This is because most Internet survey platforms only initiate branching upon completion of all the questions on a page. Thus, if there are multiple questions on the page, branching off of different questions on the page isn't usually available (Peytchev et al 2006). The obvious solution to the problem is to have each question where branching is needed be desired on a separate page. However, as noted above, having related questions grouped together results in higher item-total correlation and reduces the total time needed to complete the Internet survey (Couper et al 2001).

A second issue important in designing Internet surveys is the availability of piping. This is technology built into some Internet survey platforms that permits taking a response to a question and inserting it into the next question. For example, a question could ask "which style of car do you prefer," with the options that might include: coupe, sedan, convertible, SUV, etc. The respondent says "coupe." The next question asks "Why do you prefer a coupe? (i.e., the response "coupe" from the previous question inserted in the subsequent question). Internet survey platforms that do not permit piping would necessitate a less-artful follow-up question along the lines of "...why do you prefer the style car mentions above?" or something like it. Clearly, piping, while not necessary, makes for a "cleaner" questionnaire and is consistent with studies that have shown that design factors influence response rates (Claycomb, et al 2000 ).

## Images

The fifth Internet survey design issue is the extent to which images such as ads, pictures of packages, or websites, can be inserted into the survey and what limitations exist on their inclusion. Some authors say that graphics should be used sparingly. For example, Schonlau et al (2002) note that "...graphics can significantly slow the downloading of a web page, especially if users are likely to connect to the Internet using a modem as opposed to broadband, DSL, or some other high-speed connection." Similarly, Dillman et al (1998) found that a "plain survey" (i.e., without graphics) resulted in slightly higher response rates than a "fancy survey." On the other hand, numerous situations exist where the purpose of the survey is to test a graphic image. Examples include assessing consumer preference for or confusion with claims in advertising copy or on a package.

When considering inclusion of an image, four factors need to be considered. First, what limit exists on the Internet survey platform as to the size of image that can be included? Second, to what extent does the inclusion of an image slow the download of the survey and/or the responses? Third, can the image be scrolled on the Internet survey platform or is it static? Fourth, can questions related to the image included with the image (e.g., an open-ended question about the meaning of an ad or claim in an ad)?

If the inclusion of the image slows the download of the survey, it is likely to have a significant impact on completion rates (Couper, et al 2001). If the limit on image size is too low, a bias may be introduced since the image may not approximate what the consumer sees in the marketplace. If, on the other hand, the image is static

answering the same way on all scaled questions) and racing through the questionnaire (Flicker et al 2005). Strategies that can be used to improve the quality include employing "traps" and/or "speed bumps."

"Traps" are pass/fail directions that assure the respondent is reading the questions carefully (Balden 2008). For example, in the middle of the survey the respondent is directed to give a particular response, e.g., "mark 'somewhat disagree' on the following scale." If the respondent doesn't respond correctly, they are immediately skipped to the end and dropped from the survey. Another vehicle to enhance the quality of the data that can be used is a "speed trap." Here the researcher determines a range of time to complete a series of scaled questions and, if the Internet survey platform permits it, builds a timer for the question into the survey. If the respondent takes less time than the "minimum likely time to complete" he/she is dropped from the study. A third "trap" that can be used is an analysis of the overall time taken to complete an Internet survey. Since many Internet survey platforms include the time it took to complete the survey as part of the resulting data, speed traps can also be used, ex-post facto, to determine whether the respondent completed the survey in a reasonable period or, alternatively, so quickly that the quality of the data is questionable.

"Speed bumps" are Internet survey design methods which are intended to slow the respondent down during the survey. Speed bumps include reversing response options or scales, using questions with open-ended responses, and having questions with mandatory responses. In addition, ex-post facto analysis of responses can be used to determine whether a particular respondent "straight-lined" responses (Flicker et al 2005).

## Issues Related to Results

The final issues related to Internet surveys relate to the extent to which the researcher and/or other stakeholders have access to the data both in real-time, i.e., during the survey and in a research-accessible format. As noted above, most Internet survey platforms permit real-time tracking of results, either as each question is submitted (i.e., the interactive survey design) or upon completion of the whole survey (i.e., the scrolling survey design). Some Internet firms also permit other stakeholders, i.e., colleagues or clients, to track the data. If that option isn't available with a particular Internet firm, then the only way stakeholders can share the data in real-time is by being provided access to the researcher's confidential ID and passwords. This, of course, may compromise the confidentiality of other studies.

Finally, a critical component in any Internet survey platform is the ability to download the data for analysis with commonly used software packages. Most Internet survey platforms provide limited cross-tabulations through "filters" placed on particular questions. Many also permit downloading of raw data into "comma delimited" or ASCII format, thus permitting analysis using powerful software programs such as SAS and SPSS.

# CONCLUSION

While there are a number of factors to consider when deciding on an Internet survey platform, the summary presented here provides a guide that can be used by academic researchers as well as experts considering an Internet survey. What is obvious is that Internet surveys are here to stay and will represent an ever increasing percent of survey research in the Twenty-first Century. Moreover, as Internet survey technology increases, more creative ways to utilize this research will be developed, only adding to the power of this important new tool available to researchers. Thus, from a managerial perspective, additional opportunities will develop for researchers and experts to assist managers in assessing attitudes, opinions, and desires of their organizations' stakeholders and to help managers more effectively tailor their products and services to meet the needs of those individuals.

# REFERENCES

AAPOR (2000) "Interviewer Falification in Survey Research," American Association for Public Research. Accessed at: www.aapor.org/standards.

Gelb, Gabriel and Betsy Gelb (2006) "Internet Surveys for Trademark Litigation: Ready or Not, Here They Come." Trademark Reporter 97, pg. 1073-1088.

Healey, Benjamin, Terry Macpherson and Bart Kuijten (2005) "An Empirical Evaluation of Three Web Survey Design Principles," Marketing Bulletin 16 (2).

Hesser, Michael (2008) "A Comparison of Phone and Online Research: Each has it Strengths and Weaknesses" Quirks Marketing Research Review, Pg. 26.

Kaplowitz, Michael D, Timothy Hadlock, and Ralph Levine (2004) "A Comparison of Web and Mail Survey Response Rates" Public Opinion Quarterly 68 (1), Pg. 94-102.

Kittleson, M. (1995). "An Assessment of the Response Rate via the Postal Service and e-Mail." Health Values 18 (3), pg. 27-29.

Manfreda, Lozar and Vasja Vehovar (2004) "A Meta-Analysis of Response Rates in Web Surveys Compared to Other Survey Modes," Proceedings: American Association for Public Opinion Research.

McCarthy, J. Thomas (2006). McCarthy on Trademarks and Unfair Competition (4th Ed), 32:157ff

Nielsen, Jakob and Hoa Loranger (2006??) Prioritizing Web Usability

Pew Research (2007). "Pew Internet & American Life Project Surveys" Accessed at: PewInternet.org/trends/Internet_Activities.

Peytchev, Andy, Mick P. Couper, Sean McCabe, Scott Crawford (2006) "Web Survey Design: Paging Versus Scrolling," Public Opinion Quarterly 70 (4), pg.596-608.

Porter, Stephen R. and Michael E. Whitcomb (2003) "The Impact of Contact Type on Web Survey Response Rates." Public Opinion Quarterly 67(4), pg 579-588.

Ray, N. K. Griggs, and S. Tabor (2001) "Web Based Survey Research Workshop." WDSL. Accessed at://telecomm.boisestate.edu/research.

Roster, Catherine A, Robert D. Rodgers, Gerald Albaum, and Darin Klein (2004). "A Comparison of Response Characteristics from Web and Telephone Surveys." International Journal of Marketing Research 46 (3) pg 359-373.

Shaefer, David R and Don A. Dillman (1998) "Development of a Standard E-Mail Methodology" Public Opinion Quarterly 62 pg. 378-397.

Shonlau, Matthias, Ronald Fricker, and Marc Elliott (2002) "Conducting Research via E-Mail and the Web" Monograph Rand Corporation. Accessed at: www.rand.org/pubs/monograph_reports/MR1480.

Swann, Jerre B. "A 'Reading' Test or a 'Memory' Test: Which Survey Methodology is Correct? The Trademark Reporter 95, pg. 876-881.

Zang, Y (1999) "Using the Internet for Survey Research: A Case Study." Journal of the American Society for Information Science, 51(1), pg. 57-68.