

---

# Methods

---

## Telephone and Web: Mixed-Mode Challenge

*Jessica Greene, Howard Speizer, and Wyndy Wuitala*

---

**Objective.** To explore the response rate benefits and data limitations of mixing telephone and web survey modes in a health-related research study.

**Data Sources/Study Setting.** We conducted a survey of salaried employees from one large employer in the manufacturing sector in the summer of 2005.

**Study Design.** We randomized 751 subjects, all of whom had previously completed a web survey, to complete a health-related survey either by telephone (with web follow-up for nonrespondents) or over the web (with telephone follow-up).

**Principal Findings.** Survey response rates were significantly higher for the mixed mode survey than they would have been if we fielded either an exclusively telephone or web survey (25 and 12 percent higher, respectively). Telephone respondents were more likely to provide socially desirable responses to personal lifestyle questions and web respondents provided more missing data when questions were difficult or did not relate to their circumstances. Telephone respondents were also more likely to “agree” to knowledge statements and provide the same response across a series of items with similar response scales than were web respondents.

**Conclusions.** Mixed mode telephone/web surveys can substantially boost response rates over single-mode surveys. Modal differences in responses can be minimized by handling missing data options consistently in the two modes, avoiding agree/disagree formats, and not including questions on personal lifestyle or other sensitive issues.

**Key Words.** Mixed mode survey, survey mode effects, web survey, telephone survey

---

The challenges of conducting telephone-based survey research have increased substantially over the last decade. The emergence and popularity of caller id, call blocking, answering machines, and cellular phones have contributed to falling survey response rates and higher administration costs (Link and Oldendic 1999; Steeh et al. 2001; Tourangeau 2004). The reduction in response rate to the CDC’s telephone-based Behavioral Risk Factor Surveillance System (BRFSS) is illustrative of this trend. Between 1996 and 2001 the BRFSS

experienced a 12 percentage point drop in response rate, from 63 to 51 percent (Link and Mokdad 2005b).

Over the same time period, the web has become a viable option for conducting surveys. The promises and challenges of web survey research have been well documented (Taylor 2000; Couper 2000; Fricker and Schonlau 2002; Fricker et al. 2005). Web surveys are generally less expensive, faster to administer, and can display more complex visual information than other survey modes. But the challenges are also notable. Over a quarter of adults in the United States still do not have access to the Internet, and those without access are disproportionately poor and have low levels of educational achievement (Madden 2006). There is also no e-mail address analog to the national telephone sampling frame and some have documented low response rates to web surveys (Fricker and Schonlau 2002). Finally, people may respond differently to questions when they are posed on the web versus over the telephone.

This last issue—potential differences in responses to questions posed in web and telephone surveys—is the focus of this study. Understanding the extent to which survey mode may influence responses is important, particularly given the emergence of research that mixes telephone and web data collections (Olmstead, White, and Sindelar 2004; Satia, Galanko, and Rimer 2005; Greene et al. 2006). Researchers employ mixed mode surveys to boost response rates and reduce survey costs. Typically, researchers begin with one mode, usually the less expensive, and follow-up with nonresponders using a more costly mode (de Leeuw 2005). Some offer two modes concurrently or, in longitudinal studies, switch modes from one wave of data collection to another (Dillman and Christian 2005). The response rate and cost benefits of telephone/web-mixed mode surveys should be considered in light of differences in response patterns that are caused by survey mode. To date, however, these differences are not well understood.

There is a body of literature on differences in responses to mail and telephone surveys (Dillman et al. 1996; Dillman 2000; Tourangeau 2004). Dillman et al. (1996) outline three main differences between these two modes. The first involves the presence or absence of an interviewer. Over the telephone, the interviewer creates a social interaction that often results in respondents casting themselves in a more favorable light, particularly for questions

---

Address correspondence to Jessica Greene, Ph.D., Assistant Professor, Department of Planning, Public Policy and Management, 1209 University of Oregon, Eugene, OR 97403. Howard Speizer, M.B.A., is with RTI International, Chicago, IL. Wyndy Wiitala, Ph.D., is with Market Strategies Inc., Livonia, MI.

on sensitive topics (Dillman et al. 1996). The impact of the interviewer has been confirmed in studies that find that telephone respondents give more socially desirable responses to an interviewer than to a computer (Turner et al. 1998, 2005). The presence of an interviewer is also hypothesized to result in respondents more frequently agreeing to the premise of a question or acquiescing due to socialization toward being agreeable or deference to the interviewer. Empirical evidence, however, is mixed over whether or not there are modal differences in acquiescing (Bowling 2005).

Second, Dillman et al. highlight the difference between the visual presentation of survey questions in mail surveys and the aural communication required in telephone surveys. Mail respondents are able to review questions and the corresponding response categories, while telephone respondents typically hear the information read to them once. The cognitive burden required to process the survey question, formulate an answer, and identify the most appropriate answer category is substantially greater for telephone respondents. One manifestation of this difference is that telephone respondents more often choose the last of several options offered, compared with mail respondents, who are more likely to select the first option listed (Dillman et al. 1995).

Dillman et al. detail a third key difference between mail and telephone surveys: who controls the pace and order of the interview. Dillman et al. (1996) hypothesize that mail respondents ability to pace themselves, examine the full survey, and change prior answers would result in their providing more consistent and even-handed answers than telephone respondents, but the empirical evidence remains inconclusive.

In these respects, web surveys are quite similar to mail surveys: neither require an interviewer; both rely upon visual communication; and respondents are able to pace themselves. Web surveys also differ from mail surveys in some important ways. Web surveys can use skip patterns, offer tailored graphic displays, restrict answers to questions, and control what prior answers can be reviewed and changed. Given the similarities between the two survey modes, however, it is likely many of the differences observed between mail and telephone surveys exist between web and telephone surveys.

To date, only a small body of published research has compared modal differences in survey responses between the telephone and the web. Fricker et al. (2005) published the most rigorous comparison study to date, in which respondents with web access were randomized to either the web or the telephone for a science-related survey. They found no difference between web and telephone responses to attitude questions that were not sensitive, nor did

they find modal differences in acquiescence. Web respondents were, however, more likely to perform better on science knowledge questions. The authors suggest this may be due to web respondents' ability to reread questions and answer at their own pace. The survey had different missing data protocols for web and telephone respondents that influenced missing data provision. Web respondents who left a question blank were prompted to enter a response, which resulted in less missing data than for telephone respondents, who were not prompted. Another experimental study, which is unpublished, also confirms the lack of modal difference in responding to nonsensitive attitudinal questions (Speizer et al. 2005).

Other studies have examined differences in web and telephone survey responses, but have been limited by large sociodemographic differences between the web and telephone samples. As a result, disaggregating the influence of selection versus mode has been a key limitation. For instance, respondents to the web-based BRFSS were substantially more likely to report binge drinking than were telephone respondents, but they were also much younger and more educated than telephone respondents (Link and Mokdad 2005a, b). Some studies have used the statistical techniques of propensity score matching or weighting to equate the web and telephone samples on observed characteristics (Taylor 2000; Chang and Krosnick 2003; Schonlau et al. 2004). These studies have found strong evidence of social desirability bias among telephone responses for sensitive questions. Yet, the observed differences may be caused by uncontrolled for sociodemographic differences between the two groups, rather than modal differences. In sum, the influence of the web versus telephone on survey data is still unclear, and it has not been explored in depth for health-related questions.

In this study, we seek to examine the benefits and challenges of conducting a mixed web and telephone survey on health issues. We conducted an experiment randomizing white collar employees from one large employer, all of whom had previously completed a web survey, to complete a survey either by web (with telephone backup for nonrespondents) or by telephone (with web backup for nonrespondents). Using this experimental design, we examined two key research questions: (1) What is the benefit of offering two sequential modes rather than a single mode survey to response rates and the representativeness of the sample? (2) To what degree does survey mode influence responses to health-related behavior, attitude, and knowledge questions? To address the second question we test the degree to which we observe modal differences in social desirability, missing data, acquiescence, and question nondifferentiation.

## METHODS

### *Prior Mixed Mode Survey*

Our experimental study is part of a larger study on the influence of health insurance plan design. For the larger study, we conducted a web/telephone mixed mode baseline survey of employees at a large company during the summer of 2004. We opted for a mixed mode design over an exclusively telephone survey because of its lower cost and the potential for a higher response rate. The web was the primary survey mode we used and we conducted follow-up with nonrespondents first by e-mail (only for salaried employees, we had no e-mail addresses for hourly workers) and then by telephone (for all employees).

Among salaried employees in the 2004 study, we obtained a significantly higher response rate from the mixed mode approach than we would have if we had fielded an exclusively web survey (76 versus 63 percent). We also observed differences in responses by survey mode on some items. Telephone respondents reported being in better health and having higher levels of patient activation than web respondents, for example. Given our design, it was unclear whether these differences were due to survey mode or to selection of who completed the survey on the web versus telephone. To distinguish between these possibilities, we undertook the following experiment.

### *Study Design*

We designed the experiment into the second wave of data collection (summer 2005). We identified the 751 salaried employees who completed the baseline survey on the web and were still employed by the company. We randomly selected 250 of them to complete the second wave survey over the telephone (this number balanced cost and analytic power). After at least 12 attempts by telephone, nonrespondents were given the opportunity to complete the survey by web. The remaining 501 subjects were randomized to complete the survey over the web. Telephone follow-up was used for those who did not respond after three e-mail reminders.

The questionnaire was not modified for use in the two different survey modes. Interviewers read survey questions over the telephone and typed responses directly into the web interface. There was, however, a difference in the handling of missing data. Web respondents were able to view a missing data option for most questions. For telephone respondents, these response options were not read aloud. The impact of this difference is among the questions we examine in the study.

### *Measures*

We investigated whether survey mode influenced responses on questions about health status, health care utilization, health-related behaviors, and knowledge of the medical care delivery system.

*Health Status.* The first measure is the widely used self-report of health status, in which respondents assess their health as excellent, very good, good, fair, or poor. The second measure is an enumeration of chronic conditions. Respondents were asked whether a doctor had ever told them they had 10 specific chronic conditions (e.g., high blood pressure, arthritis, and diabetes), and we calculated a summary measure of the number of the conditions they reported.

*Health Care Utilization.* We asked respondents to report separately the number of doctor visits, hospitalizations, and emergency department visits they had in 2005. Since the survey was administered in the summer of 2005, these items required approximately 6 months of recall. Web respondents viewed a “don’t know” option for each of these questions.

*Health-Related Behaviors.* Four sets of health-related behavior measures were used in the study. First was the patient activation measure (PAM), which is a scale with strong psychometric properties that assesses the degree of knowledge, skill, and confidence respondents have in managing their health (Hibbard et al. 2004, 2005). It is comprised of 13 items that use a four-point Likert scale (agree strongly–disagree strongly). Web respondents also viewed a “doesn’t apply to me” response option. Respondents who reported no chronic health conditions answer slightly different versions of the PAM items than do those with chronic conditions (see Table 5 for items). We examine the individual items as well as a summary score, computed using Rasch analysis, ranging between 0 and 100.

The second set of health-related behavior questions focused on the frequency of engaging in healthy eating and exercise. Respondents were asked how often in a typical week (ranging from “never” to “every day”) they “exercise,” “limit how much fat is in your diet,” and “eat five or more servings of fruits or vegetables in a day.”

The third set of health behavior questions related to engaging in cost-saving risky health behaviors. Respondent were asked if they did any of the

following in 2005 in order to save money: “postpone or delay going to the doctor,” “decide not to go to the doctor when you thought you should have,” “decide against having a lab test,” “not fill a prescription,” or “take a lower dose of a prescription drug than was recommended.” In addition to “yes” and “no” options, web respondents viewed a “don’t know” option for these items.

The final set of health behavior measures related to health information seeking. Respondents were asked whether they or a household member used each of the following sources of information in 2005: their health plan website, another website for health information, a telephone nurse advice line or health coach, a website or booklet that compares prescription drug costs, a website or booklet that compares the quality of hospitals, or a website or book for handling symptoms. Again, web respondents viewed a “don’t know” option as well as “yes” and “no” options.

*Knowledge of Medical Delivery System.* We used five items developed by Newhouse, Ware, and Donald (1981) to assess respondent knowledge of the medical delivery system. The items assess general understanding of physician hospital privileges, variations in the way physicians practice medicine, physician licensing, and the limits of medicine. Respondents were asked whether they agreed or disagreed with each statement (items with correct answers are in Table 6). Web respondents were provided with a “don’t know” option as well.

### *Analysis*

To answer the first research question—what is the benefit of offering sequential mixed mode design over a unimodal design, we calculated four response rates. For those randomized to the telephone, we computed the response rate for those who responded only by telephone as well as the response rate for those who responded either by telephone or web. Similarly, for those randomized to the web, we computed rates for web only responders and all responders.

We then explored whether the additional respondents generated by offering the second mode differed with respect to demographics or health characteristics from those who responded using the primary mode. To the extent that the additional respondents differed, it suggests that the mixed mode survey reached a more representative sample. We compared 102 respondents who switched from the web to complete the survey by telephone with 168 telephone respondents randomized to the telephone who completed

by telephone. The former group of switchers included 46 people randomized to the web in the experiment who completed the survey by telephone as well as 56 people ineligible for the experiment because they responded by telephone in both waves of the survey. Given our interest in modal differences in survey responses, we sought to compare characteristics of switchers and non-switchers using data provided in one consistent mode. This was only possible for telephone respondents due to the small number of people who switched from telephone to the web.

To answer the second research question on modal differences in survey responses, we focused only on those respondents in the experimental study who completed the survey in the intended mode. That is, we excluded switchers. In examining the two experimental groups' sociodemographic characteristics we used responses from the baseline survey, when all responses were provided by web, to make certain that modal differences did not influence report of sociodemographic characteristics.

Next, using descriptive statistics we examined whether telephone and web respondents provided different responses to the health-related measures. Specifically, we were interested in whether there were differences in social desirability and missing data. We also tested whether mode influenced the extent of acquiescence. The knowledge questions were in an "agree"/"disagree" format, which enabled us to examine acquiescence, as distinct from "agreeing" to socially desirable questions. Finally, we examined nondifferentiation, or the degree to which respondents offer the same response option across a series of items. The nondifferentiation analysis was conducted using the PAM items, which are a good test for nondifferentiation because RASCH analysis has demonstrated that the items vary in difficulty. We used a measure of nondifferentiation developed by Mulligan et al. (2001), and calculated the likelihood of selecting the same response for all items, or straight lining.

## RESULTS

Offering two survey modes sequentially significantly increased survey response rates compared with offering either an exclusively telephone or web-based survey. Sixty-seven percent of those randomized to the telephone group completed the survey over the telephone (168 out of 250). When we included the 41 people assigned to the telephone who switched mode to complete the survey over the web, the response rate for this group increased to 84 percent (209 out of 250). Similarly, 75 percent of those randomized to the web



completed the survey over the web (377 out of 501). Including the 46 switchers, who completed the survey by telephone, the response rate also rose to 84 percent (423 out of 501).

We tested whether those who switched survey mode (web to telephone) differed from those completing the survey in the assigned telephone mode. To the extent that the switchers differed in terms of sociodemographics, it suggests that including the telephone follow-up mode was successful in reaching a broader group of respondents and, thus, a more representative sample. We found that those who switched to the telephone were twice as likely to be of minority race/ethnicity than nonswitcher telephone respondents (19 versus 9 percent,  $p < .05$ ). The switchers also reported less financial security than nonswitchers. We did not observe significant differences on other sociodemographic measures (e.g., gender, age, education, and health status). In sum, offering the mixed mode design increased the response rate by 9–17 percentage points, and appears to have included more minority and low-income respondents.

We now turn to our analysis of modal differences in survey responses. We compare responses between the two experimental groups (i.e., those who completed the survey in the assigned mode). Table 1 shows that there were no statistically significant sociodemographic differences between the web and telephone experiment respondents. The randomization appears to have been effective in creating demographically similar groups. Overall, the study participants are generally affluent, well educated, and healthy.

Table 2 presents comparisons between web and telephone experiment respondents for health status and utilization measures. Like most prior reports (Hochstim 1967; McHorney, Kosinski, and Je 1994), we find that respondents reporting self-rated health to an interviewer (over the telephone) are significantly more likely to report better health than when self-reporting (on the web). Telephone respondents were almost 50 percent more likely to report excellent health compared with web respondents (20 versus 14 percent). We can fully attribute this difference to mode because in the prior year, when all responses were over the web, there were no differences in health status between these two groups. For measures of health care utilization and number of chronic conditions we find no significant differences by mode. Nor was there any more missing data for web respondents, despite their viewing a “don’t know” option.

We also observe substantial modal differences for personal lifestyle questions that related to healthy eating and exercising (Table 3). Phone responders compared with web respondents were almost three times as likely to

Table 1: Characteristics of Respondents by Experiment Status as Reported in Prior Year's Web Survey

<i>Employee Characteristics</i>	<i>Percent of Total Sample (n = 545)</i>	<i>Experiment Groups</i>		<i>p-Value</i>
		<i>Web % (n = 377)</i>	<i>Telephone % (n = 168)</i>	
Gender				
Male	60.2	62.6	54.8	.084
Female	39.8	37.4	45.2	
Age				
< 40	20.9	20.2	22.6	.745
40–50	40.2	40.1	40.5	
50–60	38.9	39.8	36.9	
Race/ethnicity				
Non-Hispanic white	89.9	89.5	90.7	.577
Hispanic or nonwhite	10.1	10.5	9.3	
Marital status				
Single	12.9	13.7	11.3	.654
Married/living with partner	76.9	76.7	77.4	
Other	10.1	9.6	11.3	
Education				
High school graduate	9.5	9.3	10.1	.316
Some college	35.5	33.8	39.3	
College graduate	32.3	32.2	32.7	
Graduate studies	22.6	24.8	17.9	
Family income				
< \$50,000	14.3	13.8	15.5	.177
\$50,000–\$75,499	26.4	26.0	27.4	
\$75,000+	42.0	40.6	45.2	
Prefer not to report	17.2	19.6	11.9	
2004 Self-rated health				
Excellent	19.8	18.7	22.2	.436
Very good	37.5	39.6	32.9	
Good	35.5	34.2	38.3	
Fair/poor	7.2	7.5	6.6	

report eating five fruits or vegetables daily (16 versus 6 percent) and were twice as likely to report daily exercising (20 versus 9 percent). We tested whether collapsing response options would remedy the strong mode effect, but the effect was still pronounced when we ran statistical testing with dichotomized categories.

Table 4 displays the risky cost saving and information seeking behaviors by mode. Telephone respondents tended to give more socially desirable answers; however, the modal differences were relatively small (generally within

Table 2: Comparison of Responses to Health-Related Questions, by Experiment Status

<i>Measure</i>	<i>Health-Related Measures</i>		<i>Mean Number of Items with Missing Responses</i>	
	<i>Web</i> ( <i>n</i> = 367)	<i>Telephone</i> ( <i>n</i> = 168)	<i>Web</i> ( <i>n</i> = 367)	<i>Telephone</i> ( <i>n</i> = 168)
<b>Health status</b>				
Self-rated health (%)				
Excellent	14.3	20.2**	n/a	n/a
Very good	34.5	44.0		
Good	41.1	30.4		
Fair/poor	10.1	5.4		
Mean number of chronic conditions (0–10 range)	1.1	1.1	0.1	0.0
<b>Health care utilization in 2005</b>				
Mean number of doctor visits	1.9	2.2	0.0	0.0
Mean number of hospitalizations	0.0	0.2	0.0	0.0
Mean number of ED visits	0.1	0.1	0.0	0.0

\*\* $p < .01$ .

three percentage points) and only in one case was the difference significant. There were little missing data for these items (never > 2 percent), and it did not differ by mode.

Telephone respondents also had slightly higher summary patient activation levels (68 versus 66 for web respondents), but again the difference was

Table 3: Responses to Healthy Behavior Items, by Experiment Status

<i>How Often Do Recipients Do Each of the Following in a Typical Week</i>	<i>Never (%)</i>	<i>Some Days (%)</i>	<i>Most Days (%)</i>	<i>Every Day (%)</i>
<b>Limit how much fat is in your diet</b>				
Web	5.3	37.4	49.1	8.2***
Phone	9.5	24.4	48.2	17.9
<b>Exercise</b>				
Web	4.5	50.7	36.3	8.5**
Phone	5.4	38.1	36.3	20.2
<b>Eat five or more servings of fruits or vegetables in a day</b>				
Web	17.0	52.5	24.7	5.8
Phone	5.4	50.6	28.6	15.5***

\*\* $p < .01$ ; \*\*\* $p < .001$ . (Note: respondents were not offered missing data options for these items in either mode.)

not significant. Table 5 details the modal differences for the specific patient activation items. In four cases the differences were statistically significant. In each of these four cases, telephone respondents were more likely to “agree” or “strongly agree” to socially desirable statements than were web respondents. Web respondents, who viewed a missing data option, provided significantly

Table 4: Responses to Risky Cost Saving Behaviors and Information Seeking Items, by Experiment Status

<i>Behaviors in 2005</i>	<i>Yes (%)</i>	<i>No (%)</i>
Risky cost saving behaviors		
Postpone or delay going to the doctor to save money		
Web	19.9	80.1
Phone	16.1	83.9
Decide not to go to the doctor when you thought you should have to save money		
Web	16.1	83.9
Phone	13.1	86.9
Decide against having a lab test to save money		
Web	8.7	91.3
Phone	7.2	92.8
Not fill a prescription to save money		
Web	6.7	93.3
Phone	6.8	93.2
Take a lower dose of a prescription drug than was recommended in order to save money		
Web	7.5	92.5
Phone	4.5	95.5
Information seeking		
Use health plan’s website		
Web	50.4	49.6
Phone	53.0	47.0
Use any website for health information, other than health plan’s		
Web	58.2	41.8
Phone	65.5	34.5
Use a telephone nurse advice line or health coach		
Web	13.8	86.2
Phone	15.5	84.5
Use a website or booklet that compares prescription drug costs		
Web	11.4	88.6*
Phone	18.5	81.5
Use a website or booklet that compares the quality of hospitals		
Web	8.5	91.5
Phone	13.7	86.3
Use a website or book that helps you figure out how to handle a symptom or health problem		
Web	48.8	51.2
Phone	53.0	47.0

\* $p < .05$ . (There was <2% missing data for these items, and as it did not differ by mode, it is not reported in the table.)

Table 5: Responses to the Patient Activation Measure's 13 Items, by Experiment Status

<i>PAM Items</i>	<i>Disagree and Strongly Disagree (%)</i>	<i>Agree (%)</i>	<i>Strongly Agree (%)</i>	<i>Doesn't Apply to Me (%)</i>
a. Taking an active role in my own health care is the most important factor in determining my health and ability to function				
Web	4.2	51.5	43.2	1.1
Phone	3.0	50.6	45.2	1.2
b. I know how to prevent (further) problems with my health (condition)				
Web	8.0	64.7	22.8	4.5
Phone	7.1	67.9	22.0	3.0
c. I know what each of my prescribed medications does				
Web	2.9	37.7	29.7	29.7*
Phone	0.0	43.2	36.9	20.8
d. I understand the nature and causes of my health (condition/problems)				
Web	10.1	50.7	22.3	17.0
Phone	6.0	56.5	23.8	13.7
e. I am confident that I can follow through on medical treatments I (may) need to do at home				
Web	2.4	48.5	34.7	14.3**
Phone	1.8	50.0	44.6	3.6
f. I am confident that I can take actions that will help prevent or minimize some symptoms or problems associated with my health (condition)				
Web	4.2	57.0	32.4	6.4*
Phone	1.8	58.3	38.1	1.8
g. I have been able to maintain the lifestyle changes for my health (condition) that I have made				
Web	10.1	55.7	22.0	12.2
Phone	5.4	59.5	26.8	8.3
h. I am confident I can figure out solutions when new situations or problems arise with my health (condition)				
Web	16.2	60.2	19.9	3.7
Phone	10.1	64.3	20.2	5.4
i. I am confident I can tell a doctor concerns I have even when he or she does not ask				
Web	3.7	53.6	41.6	1.1
Phone	2.4	48.8	48.8	0.0
j. I am confident that I can tell when I need to get medical care and when I can handle a health problem myself				
Web	6.9	58.9	33.2	1.1
Phone	3.6	58.3	37.5	0.6
k. I know the different medical treatment options available for my health condition				
Web	13.8	47.7	15.4	23.1
Phone	9.5	57.7	15.5	17.3
l. I am confident that I can maintain lifestyle changes, like diet and exercise, even during times of stress				
Web	20.4	55.2	21.2	3.2*
Phone	13.7	58.9	26.8	0.6

Table 5. *Continued*

<i>PAM Items</i>	<i>Disagree and Strongly Disagree (%)</i>	<i>Agree (%)</i>	<i>Strongly Agree (%)</i>	<i>Doesn't Apply to Me (%)</i>
m. When all is said and done, I am the person who is responsible for managing my health (condition)				
Web	2.9	46.2	50.7	0.3
Phone	1.8	42.3	55.4	0.6

\* $p < .05$ ; \*\* $p < .01$ . (There are two versions of the PAM, one for those with chronic conditions and one for those without. The brackets indicate the slight differences in wording in the two versions.) Patient Activation Measure Copyright 2005–2007, University of Oregon, All Rights Reserved.

more missing data than telephone respondents in three out of the four cases (items c, e, f; analysis not shown). These three items assume that respondents have a health problem (e.g., “I know what each of my prescribed medications does”). Items that were applicable regardless of health status, like “I am confident I can tell a doctor concerns I have even when he or she does not ask,” had little missing data and no modal differences.

We also tested whether there were modal differences in nondifferentiation, or providing the same response option across the patient activation items. Contrary to Fricker et al. (2005), we found that web respondents provided more differentiated responses than telephone respondents. Phone respondents were twice as likely as web respondents to give the exact same response across all 13 items or “straight line” (16 versus 8 percent,  $p < .01$ ), and they had higher value on the Mulligan et al. (2001) measure (0.6 versus 0.5,  $p < .01$  on a scale in which “0” indicates all responses are different and “1” indicates all the responses are the same).

Our final analysis compared responses to knowledge questions by mode (Table 6). Web respondents who viewed a “don’t know” category were substantially more likely to pick this response than were telephone respondents who had to volunteer a “don’t know” response. Equal percentages of web and telephone respondents correctly answered the two items for which “disagree” was the correct response (items b and c). For these items, web respondents were more likely to select “don’t know” while telephone respondents were more than twice as likely to agree (or acquiesce) to the statements. Greater acquiescence among telephone respondents is also likely the reason that they more often correctly answered items that were in fact true (a, d, and e). It seems highly unlikely that web and telephone respondents were equally knowledgeable for items b and c, which were false statements, but that telephone

Table 6: Responses to the Knowledge Items, by Experiment Status

<i>Knowledge Items</i>	<i>Agree (%)</i>	<i>Disagree (%)</i>	<i>Do Not Know (%)</i>
a. Some operations done by surgeons are not really necessary			
Web	<b>61.3</b>	11.7	27.1***
Phone	<b>72.6</b>	16.7	10.7
b. If you have to go into the hospital, your doctor can get you admitted to any hospital you prefer			
Web	25.5	<b>28.4</b>	46.2***
Phone	51.2	<b>29.8</b>	19.0
c. Doctors are checked every few years, before their licenses are renewed			
Web	13.3	<b>22.5</b>	64.2***
Phone	33.9	<b>24.4</b>	41.7
d. For many illnesses, doctors just do not have any cure			
Web	<b>61.3</b>	17.0	21.8***
Phone	<b>76.8</b>	19.0	4.2
e. Two doctors who are equally good at their jobs may still suggest very different ways of treating the same illness			
Web	<b>89.1</b>	2.4	8.5*
Phone	<b>95.8</b>	2.4	1.8

\* $p < .05$ ; \*\*\* $p < .001$ . (Bolded answers are correct.)

respondents were more knowledgeable for the three true items. When we computed a summary measure of knowledge performance, telephone respondents performed significantly better than web respondents (mean scores of 3.0 versus 2.6, respectively,  $p = .001$ ) as overall there were more true items than false items.

## DISCUSSION

This study uses a randomized experiment to examine the benefits and limitations of mixed mode web and telephone surveys. We find that the mixed mode design increased response rates over fielding exclusively telephone or web-based surveys by 25 and 12 percent, respectively. Our findings further demonstrate that the higher response rate resulted in a more representative sample, including more ethnically diverse and low-income respondents. The study also finds that survey mode influences social desirability bias, missing data, acquiescence, and nondifferentiation—often in nontrivial ways. Below we discuss these patterns and discuss the practical implications for designing mixed mode surveys.

The modal differences in social desirability bias that we observed varied in magnitude considerably, depending on the type of questions. For objective events such as reports of physician diagnosis of chronic conditions and health care utilization there were no modal differences observed. For other items in the survey with socially desirable answers (information seeking, risky cost saving behaviors, and managing one's health) we observed a trend toward more socially desirable answers among telephone respondents, but the difference was slight and not generally statistically significant. Where we observed the greatest modal differences were for lifestyle issues of healthy eating and exercising (including one patient activation item mentioning diet and exercise) and health status. The magnitude of modal differences was quite substantial for these items, suggesting that these types of items not be included in mixed mode surveys or that analyses combining web and telephone data statistically adjust for survey mode.

Our results also revealed differences in provision of missing data for specific types of questions. Despite the fact that web respondents viewed a missing data option that was not mentioned over the telephone, there were very low levels of missing data regardless of mode for most questions. There were, however, two circumstances in which web respondents provided substantially more missing data: answering relatively difficult knowledge questions and questions that assumed a health problem. In both circumstances, web respondents' higher levels of missing data appears to be appropriate and not due to satisficing or taking cognitive "short cuts." These observed modal differences underscore the importance of developing mixed mode surveys that are exactly the same for each mode, including the treatment of missing data.

We also found modal differences in agreeing to factual statements assessing knowledge. Telephone respondents were more likely to "agree" than were web respondents. Consequently, when the correct answer to a question is "agree," telephone respondents will likely outperform web respondents. Based on these findings, we recommend avoiding agree/disagree and true/false type questions in mixed mode surveys.

Our findings suggest that researchers can reduce modal differences by designing surveys that can be fielded identically in the two modes, avoiding sensitive questions about lifestyle, and not including true/false or agree/disagree questions. For surveys on personal lifestyle issues or sensitive matters, we recommend not mixing telephone and web modes. If non-web follow-up is necessary for such a survey, researchers should consider telephone audio computer-assisted self-interviewing or other modes that do not involve an



interviewer. Finally, given the substantial modal differences we observed for some items, we caution researchers against mixing modes across waves of longitudinal studies. While the lower cost of the web may incentivize researchers to convert telephone respondents from one wave of a study to web respondents in subsequent waves, the change in mode would confound isolating changes due to the factor of interest.

In addition to the lessons for mixed mode survey design, this study highlights the comparative strengths of web surveys. For our Internet savvy sample, we find that the web results in less social desirability bias, acquiescence, and nondifferentiation.

The study findings should be interpreted in light of several limitations. The primary limitation relates to generalizability. Our sample is relatively well educated and was selected based upon having responded to a prior survey on the web. More research is needed to examine the effect of mode for those with lower socioeconomic status and less familiarity with the web. Additionally, we only examine one new survey mode, the web, while a number of other emerging survey approaches are being used. It will be important for survey researchers going forward to conduct randomized tests of the differences in responses when introducing new technologies.

## ACKNOWLEDGMENT

The authors would like to thank “The Changes in Health Care Financing and Organization” (HCFO) initiative, a program of The Robert Wood Johnson Foundation, for providing support for this research.

## REFERENCES

- Bowling, A. 2005. “Mode of Questionnaire Administration Can Have Serious Effects on Data Quality.” *Journal of Public Health* 27 (3): 281–91.
- Chang, L., and J. Krosnick. 2003. “National Surveys via RDD Telephone Interviewing vs. the Internet: Comparing Sample Representativeness and Response Quality.” Unpublished Manuscript.
- Couper, M. 2000. “Web Surveys: A Review of Issues and Approaches.” *Public Opinion Quarterly* 64: 464–94.
- Dillman, D. 2000. *Mail and Internet Surveys*, 2d Edition. New York: John Wiley and Sons.

- Dillman, D., T. Brown, J. Carlson, E. Carpenter, F. Lorenz, R. Mason, J. Saltiel, and R. Sangster. 1995. "Effects of Category Order on Answers to Mail and Telephone Surveys." *Rural Sociology* 60: 674–87.
- Dillman, D., and L. Christian. 2005. "Survey Mode as a Source of Instability in Responses across Surveys." *Field Methods* 17 (1): 30–52.
- Dillman, D., R. Sangster, J. Tarnai, and T. Rockwood. 1996. "Understanding Differences in People's Answers to Telephone and Mail Surveys." In *Advances in Survey Research*, edited by M. Braverman and J. Slater, pp. 45–62. San Francisco: Jossey-Bass.
- Fricker, R., and M. Schonlau. 2002. "Advantages and Disadvantages of Internet Research Surveys: Evidence from the Literature." *Field Methods* 14 (4): 347–67.
- Fricker, S., M. Galesic, R. Tourangeau, and T. Yan. 2005. "An Experimental Comparison of Web and Telephone Surveys." *Public Opinion Quarterly* 69 (3): 370–92.
- Greene, J., J. H. Hibbard, A. Dixon, and M. Tusler. 2006. "Which Consumers Are Ready for Consumer-Directed Health Plans?" *Journal of Consumer Policy* 29 (3): 247–62.
- Hibbard, J. H., E. R. Mahoney, J. Stockard, and M. Tusler. 2005. "Development and Testing of a Short Form of the Patient Activation Measure." *Health Services Research* 40 (6, part1): 1918–30.
- Hibbard, J. H., J. Stockard, E. R. Mahoney, and M. Tusler. 2004. "Development of the Patient Activation Measure (PAM): Conceptualizing and Measuring Activation in Patients and Consumers." *Health Services Research* 39 (4, part 1): 1005–26.
- Hochstim, J. 1967. "A Critical Comparison of Three Strategies of Collecting Data from Households." *Journal of the American Statistical Association* 62:319: 976–89.
- Leeuw, E. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21 (2): 233–55.
- Link, M., and R. Oldendick. 1999. "Call Screening. Is it Really a Problem for Survey Research?" *Public Opinion Quarterly* 63: 577–89.
- Link, M. W., and A. H. Mokdad. 2005a. "Alternative Modes for Health Surveillance Surveys: An Experiment with Web, Mail, and Telephone." *Epidemiology* 16 (5): 701–4.
- . 2005b. "Effects of Survey Mode on Self-Reports of Adult Alcohol Consumption: A Comparison of Mail, Web and Telephone Approaches." *Journal of Studies on Alcohol* 66 (2): 239–45.
- Madden, M. 2006. *Internet Penetration and Impact*. Washington, DC: Pew Internet & American Life Project.
- Mchorney, C., M. Kosinski, and W. Je. 1994. "Comparisons of the Costs and Quality of Norms for the SF-36 Health Survey Collected by Mail versus Telephone Interview: Results from a National Survey." *Medical Care* 32 (6): 551–67.
- Mulligan, K., J. Krosnick, W. Smith, M. Green, and G. Bizer. 2001. "Nondifferentiation on Attitude Rating Scales: A Test of Survey Satisficing Theory." Manuscript under Review.
- Newhouse, J. P., J. E. Ware, and C. A. Donald. 1981. "How Sophisticated Are Consumers about the Medical Care Delivery System?" *Medical Care* 19 (3): 316–28.

- Olmstead, T., W. D. White, and J. Sindelar. 2004. "The Impact of Managed Care on Substance Abuse Treatment Services." *Health Services Research* 39 (2): 319–43.
- Satia, J. A., J. A. Galanko, and B. K. Rimer. 2005. "Methods and Strategies to Recruit African Americans into Cancer Prevention Surveillance Studies." *Cancer Epidemiology, Biomarkers and Prevention* 14 (3): 718–21.
- Schonlau, M., K. Zapert, L. P. Simon, K. H. Sanstad, S. M. Marcus, J. Adams, M. Spranca, H. Kan, R. Turner, and S. H. Berry. 2004. "A Comparison between Responses from a Propensity-Weighted Web Survey and an Identical RDD Survey." *Social Science Computer Review* 22 (1): 128–38.
- Speizer, H., K. Schneider, W. Wiitala, and R. Baker. 2005. "An Experimental Comparison between Telephone and Web." Presented at: American Association for Public Opinion Research Annual Conference, Miami, FL.
- Steeh, C., N. Kirgis, B. Cannon, and J. Dewitt. 2001. "Are They Really as Bad as They Seem? Nonresponse Rates at the End of the Twentieth Century." *Journal of Official Statistics* 17 (2): 227–47.
- Taylor, H. 2000. "Does Internet research work? Comparing electronic survey results with telephone survey." *International Journal of Market Research* 42 (1): 51–63.
- Tourangeau, R. 2004. "Survey Research and Societal Change." *Annual Review of Psychology* 55: 775–801.
- Turner, C., B. Forsyth, J. O'reilly, P. Cooley, T. Smith, S. Rogers, and H. Miller. 1998. "Automated Self-Interviewing and the Survey Measurement of Sensitive Behaviors." In *Computer Assisted Survey Information Collection*, edited by M. Cooper, R. Baker, J. Behlehem, C. Clark, J. Martin, W. Nicholls, and J. O'Reilly, pp. 455–73. New Jersey: John Wiley & Sons, Inc.
- Turner, C. F., M. A. Villarroel, S. M. Rogers, E. Eggleston, L. Ganapathi, A. M. Roman, and A. Al-Tayyib. 2005. "Reducing Bias in Telephone Survey Estimates of the Prevalence of Drug Use: A Randomized Trial of Telephone Audio-Casi." *Addiction* 100 (10): 1432–44.