Journal of Official Statistics, Vol. 22, No. 2, 2006, pp. 293–312

Can Web and Mail Survey Modes Improve Participation in an RDD-based National Health Surveillance?

Michael W. Link and Ali Mokdad[1]

Participation in most random-digit dialed (RDD) telephone surveys is declining, leaving researchers to question the quality of the data produced. As in the case of many other RDD studies, response rates for the Behavioral Risk Factor Surveillance System (BRFSS) have declined in recent years, prompting consideration of alternative modes to increase participation. In this study, web and mail versions of the BRFSS questionnaire were administered to potential respondents drawn from the standard BRFSS telephone sampling frame and reverse-matched to identify valid mailing addresses. Telephone survey follow-up was conducted with web and mail survey nonrespondents. Results were compared with those from the on-going computer-assisted telephone interviewing (CATI) BRFSS. The findings suggest that self-administered modes when used in conjunction with telephone follow-up can improve levels of participation, but may also increase differences between respondents and nonrespondents on measures of interest. As a result, overall nonresponse bias may not have been reduced despite increases in response rates.

Key words: Web surveys; mail surveys; survey bias; RDD.

1. Introduction

Willingness to participate in random-digit dialed (RDD) telephone surveys has been declining for at least the past decade (de Leeuw and de Heer 2002). The growth in telemarketing efforts, increased use of cellular telephones, advancements in caller-identification and call-screening technology, and the advent of "do not call" lists have helped to accelerate this drop in participation. Increased levels of nonresponse when combined with differences in measures of interest – such as attitudes, behaviors, and beliefs – for respondents and nonrespondents threaten the validity and reliability of data reported in probability sample surveys, potentially increasing the nonresponse bias associated with estimates made from these data (Babbie 1990; Dillman et al. 2002). Researchers have tested a variety of potential solutions, including the use of prenotification letters (Ford 1967; Dillman, Gallegos, and Frey 1976; Kennedy et al. 1998; Goldstein and Jennings 2002; Link and Mokdad 2005a), improved training of interviewers (Benson, Booman, and Clark 1951; Groves 2002; Dillman et al. 2002; Link

[1] Division of Adult and Community Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, 4770 Buford Highway, NE, Mailstop K-66, Atlanta, GA 30341-3717, U.S.A. Emails: MLink@cdc.gov and AMokdad@cdc.gov.

Acknowledgments: We thank the four state coordinators Colleen Baker (New York), Larry Shireley (North Dakota), Linda Stemnock (Indiana), and Neha Thakkar (Arkansas), as well as Ruth Jiles (Centers for Disease Control and Prevention), Janice Rush and Kerri Holloway (Clearwater Research), and Jodie Weiner and David Roe (RTI International) for assisting in developing and implementing this study.

et al. 2005), and incentives (Church 1993; James and Bolstein 1990 and 1992; Singer, Groves, and Corning 1999; Singer, Hoewyk, and Maher 2000). Use of multiple modes of questionnaire administration is another possible approach, in particular the use of web-based surveys and mail surveys as complements to telephone data collection (Dillman 1999).

Mixing survey modes has the potential for extending the reach of a survey by encouraging participation across a broader mix of the population. "The time when population access and response could usually be achieved by one mode is past," noted Dillman (2000, p. 223), "For increasing numbers of surveys, reliance on only one mode does not provide assurance of reaching or eliciting responses from most of the sampled individuals." Research has shown that some sample members prefer and respond more readily to different survey modes (Groves and Kahn 1979). For example, some sample members might like completing an interview over the telephone with the assistance of an interviewer, while others may prefer to complete a self-administered survey over the web or on paper at their own pace and with their own timing.

There are potential drawbacks, however. Different modes can, at times, produce different results even when identical questions are asked of the same persons in the sample (de Leeuw 1992; Dillman et al. 1996; Dillman 2000). In this respect, use of alternative modes may increase response rates, but also increase measurement differences. For example, studies have shown a greater likelihood of respondents giving socially desirable responses in interview surveys than in self-administered surveys (DeMaio 1984; Aquilino 1994; Dillman et al. 1996). Moreover, although a relatively large body of literature has examined multimode surveys involving combinations of telephone, mail, and face-to-face surveys, our understanding of the role of web surveys is limited (Link and Mokdad 2005b; Link and Mokdad 2005c; Schonlau, Asch, and Du 2003; McCabe et al. 2002; Couper 2000).

As the largest, state-based RDD telephone health survey in the United States, the Behavioral Risk Factor Surveillance System (BRFSS) is confronted with the issue of declining response rates and the associated questions of data reliability and validity. Use of multiple modes is one means of addressing this problem. The BRFSS is administered by the 50 state health departments as well as those in the District of Columbia, Puerto Rico, Guam, and the Virgin Islands with support from the Centers for Disease Control and Prevention (CDC). Interviews are conducted with the noninstitutionalized adult population aged 18 years or older, focusing on health practices and risk behaviors linked to chronic diseases, injuries, and preventable infectious diseases. Unfortunately, mean response rates across participating states for the BRFSS dropped 12 percentage points in a six-year period from 1996 (63 percent) to 2001 (51 percent) (Link 2003). Compounding this problem is the issue of differential participation rates among various population subgroups. As in the case of many other RDD-based telephone surveys of the adult population, BRFSS respondents tend to over-represent women, whites (non-Hispanics), older adults (particularly those age 55 and older), and those with higher levels of income and education. If BRFSS and other RDD surveys that collect similar data are to continue to meet the needs of local, state, and national researchers and policymakers for public health data, alternative means of collecting these data need to be examined.

To this end, experiments were conducted to test the effectiveness of web and mail surveys used in conjunction with telephone follow-up of nonrespondents to increase participation in BRFSS surveys. The research attempted to answer four key questions about the use of mixed-mode surveys:

1. Can alternative modes help to increase participation in the survey?
2. What types of individuals in the general public respond to web and mail surveys?
3. Can use of mixed-modes close the gap in participation among subgroups to reduce the differences between respondents and nonrespondents on measures of interest?
4. Why do some persons choose not to complete a self-administered questionnaire?

2. Methods

Two sets of experiments were conducted in four U.S. states (Arkansas, Indiana, New York, and North Dakota) during October and November, 2003. In the first experiment, letters were mailed to sample members asking them to complete the BRFSS survey via the Internet. In the second, questionnaires were mailed to sample members, which they were asked to complete and return. In both experiments, nonrespondents to these self-administered modes were followed-up by telephone interviewers beginning approximately 10 days later and asked to complete the survey by telephone. The experiments were conducted in parallel with the on-going, monthly BRFSS data collection in each of the states, providing a baseline for comparison.

2.1. Sample Selection

The multistage design allowed assessment of the effects of using multiple survey modes in various regions of the country, with different populations, and among states that varied in rates of participation in the BRFSS survey. Additionally, data collection for these states was managed by a single vendor, which provided both operational and methodological advantages. BRFSS is a state-based surveillance, allowing states to determine which organization will conduct data collection. Some state agencies have in-house data collection capabilities; others outsource the work to national or regional survey firms or universities. Because this was an initial pilot test of self-administration of the BRFSS survey, it was logistically easier and cost-effective to coordinate development and hosting of the web site and development and printing of the questionnaires for mail out with a single vendor. Methodologically this strategy ensured greater standardization in data collection procedures, programming and hosting of the web survey, and handling of mailed questionnaires across the four states. This approach also allowed the same interviewers who were conducting the regular, monthly BRFSS survey to use computer-assisted telephone interviewing (CATI) in follow-up of nonrespondents to the web or mail surveys. Use of the same interviewers eliminated the potential for bias caused by differences that can be introduced when a study is split across multiple data collection locations or vendors.

Monthly BRFSS samples for each of the five states were drawn according to standard, previously approved and CDC-monitored BRFSS protocols. The sampled telephone numbers were then reverse-matched for addresses by Marketing Systems Group using

their Listed Database, licensed from InfoU.S.A. Only households with a telephone number that could be matched with a valid mailing address were retained in the samples. The state-specific address-match rates for all sampled numbers were: New York = 31.0 percent, North Dakota = 31.8 percent, Indiana = 32.3 percent, and Arkansas = 36.5 percent. If we included only the verified and potentially eligible cases in the denominator (i.e., excluded the cases determined to be ineligible), the address-match rates ranged from 52.8 percent in New York to 69.9 percent in North Dakota.

Contact information for subjects who did not complete the self-administered questionnaire by the tenth day of the month was loaded into CATI for telephone follow-up. In the mail experiment, all nonresponding sample members were followed-up by telephone. In the web experiment, a subsampling strategy was used to select subjects for CATI follow-up, given the larger starting sample size for the web experiment which took into account the fact that not all households have access to the Internet. The number of completions by CATI follow-up for each mode was allowed to vary depending on the highest achievable response rate obtained by the end of the monthly data collection period.

2.2. *Instrument Development*

Designed by state coordinators, the BRFSS questionnaire has three parts: (1) the core component, (2) optional modules, and (3) state-added questions. The core component consists of a set of questions asked by all states without modification in order or wording. Only the core component was used in the web and mail surveys. In contrast, the baseline telephone-only data were collected using the core module and any of the optional modules or state-specific questions used by the four states as part of the regular data collection effort for the 2003 BRFSS.

Question wording, format and design, and the impact of audio versus visual stimuli are among the primary issues faced when any study is changed from a CATI design to a web or mail survey design (Dillman 2000; Couper 2000). CATI uses an interviewer-administered questionnaire, but web and mail surveys are self-administered. Wording that sounds correct in a two-person conversation on the telephone may not translate well when read from a hard-copy questionnaire or a computer screen in isolation by a survey participant. Additionally, for CATI, the layout on the screen is typically designed for the convenience of interviewers. Interviewers receive specialized training on the layout for different types of questions and how to navigate within the questionnaire. Participants using a web or mail instrument, who receive only limited, written instructions, may have greater difficulty working their way through the questionnaire.

In developing the web and mail versions of the BRFSS, we followed as closely as possible the "unimode construction" approach as described by Dillman (2000, p. 232): "the writing and presenting of questions to respondents in a way that assures receipt by respondents of a common mental stimulus, regardless of survey mode." Although the instrument was initially designed as a CATI survey, no wording changes were required in the text of the 84 core items in the BRFSS questionnaire when the survey was converted to web and mail modes.

The most significant difference between the interviewer-administered and self-administered versions was in the handling of the "don't know" and "refuse" response

categories. Studies have documented that when either of these categories is used explicitly, regardless of mode of administration, the percentage of respondents selecting "don't know" or "refuse" increases (Coombs and Coombs 1976; Converse 1976; Faulkenberry and Mason 1978). In the CATI version of BRFSS, "don't know" and "refuse" options are available for nearly every question, but interviewers are trained not to read these options and to select them only if the respondent volunteers these answers. Additionally, interviewers are trained to probe most "don't know" and "refuse" responses. Unfortunately, with self-administered modes, researchers need to make the decision up front whether to offer "don't know" and/or "refuse" as explicit options.

For these experiments "don't know" was made an explicit option on both the mail and web surveys for items that typically receive 10 percent or more "don't know" responses during regular, monthly CATI data collection. For items with less than 10 percent "don't know" responses these responses are typically ignored (declared "missing") by researchers in many analyses. At 10 percent or more, however, "don't know" is more often viewed as a substantive response and generally included in analyses. The inclusion of "don't know" as a response on items where "don't know" is more readily recorded in the CATI interview increases the comparability across the telephone, web, and mail modes. Following this reasoning, "don't know" was made explicit for four questions on the web and mail surveys. These items were pneumonia vaccination status, the respondent's race, household income, and the month and year of the last test for human immunodeficiency virus (HIV).

Unlike the "don't know" response option, which was handled identically in the mail and web versions of the questionnaire, explicit indication of "refuse" was handled differently for each survey mode. "Refuse" was not explicitly offered in either self-administered mode. Instead it was assumed that respondents to the mail survey would simply leave a question blank if they did not want to answer it. Blank responses were coded by data-entry staff as "refuse" when the mail responses were entered into the database unless the items were left blank because of a valid instruction to skip the question. In contrast, web survey respondents were provided with a "continue" button at the bottom of each screen, which they could press and the computer would record this as a "refuse" response. For several questions considered to be essential items (i.e., where the response drove major routing within the questionnaire), a pop-up box appeared to reiterate the importance of providing a response to the question and verifying that the respondent wanted to move forward without entering a response.

To develop formatting and layout of the mail questionnaire and the web survey screens, we relied primarily on previous research and the operational experience of the researchers involved in the project. The mail survey followed a design used successfully with diverse populations in previous studies, particularly those including more elderly participants (e.g., Consumer Assessment of Health Plans Survey). The questionnaire was printed as a booklet, which has been shown to be the preferred format for presenting hardcopy questionnaires (Dillman 2000). The instructions were to mark responses in the field indicated to the right of the response options. Skip logic and special instructions were clearly delineated in bold print.

The web site was designed with a home page or portal by using graphics similar to those used for the mail survey booklet. The entry portal included the study name and links to the CDC web site and four state-specific BRFSS web sites for respondents who wanted to

learn more about the study. It also contained specified fields for entering the username and password provided on the advance letter for accessing the survey. Respondents logging-on to the web site were presented with the confidentiality statement and two screens of instructions providing an overview of how to navigate within the questionnaire and how to enter responses for particular types of questions (e.g., use of radial buttons and entry of text). The questionnaire was presented with one question per screen.

2.3. Mailings

Members of households in the web and mail experiments were first notified about the study by a letter addressed to "respondent" and printed on state department of health stationery and signed by a state public health official. The letter provided a brief overview of the study and indicated the importance of having an adult in the household participate. A unique username and password were provided along with the Internet address for accessing the web site. A toll-free number was also provided for additional information.

Households in the mail experiment received a packet containing a cover letter on the same stationery and with the same signature as the letter sent to web experiment households, a copy of the questionnaire booklet, and a return envelope with postage paid. The return address on the original packet envelop was that of the state health department; the address on the return envelopes was that of the data collection vendor. Both the web letters and the mail packets were sent three days before the start of data collection, so that most potential respondents would receive them on the first of the month. The web site was made available on the day the letters were mailed for any respondent who may have received the mailing early and attempted to access the web instrument. A small percentage of respondents did access the web during this time frame.

On the fifth day of the month a reminder letter was sent to all respondents. The letters reiterated the importance of the household's participation in the survey, offered a toll-free number for questions, and again provided the user name and password for the web survey. Because the time for data collection in each month was limited, potential respondents to the mail survey received only one questionnaire, in the initial mailing.

2.4. Within Household Selection of Respondents

Respondent selection was handled differently for the CATI survey in both the baseline data collection and nonresponse follow-ups than for the self-administered surveys. For telephone interviewing, a random-selection technique (Troldahl-Carter method) was used to select a respondent within a household (Bryant 1975). Interviewers entered the number of adult males and females in the household, and the computer determined the respondent.

The decision on how to select a respondent within the household for the self-administered surveys was a point of deliberation. An automated approach similar to that used in CATI was possible with the web survey, but it was thought that the absence of an interviewer to guide the process might make this method cumbersome. To maintain comparability between the mail and web surveys, we wanted to use the same approach for both. Considering this issue from a "total error" perspective, we weighed the costs and benefits of using some form of random technique, which would maintain the random-selection properties of the sample but could lead to higher levels of nonresponse, against

the costs and benefits of allowing the household to select who would participate, which would have the opposite effect. In the end, the household members were asked to select an adult to complete the questionnaire. This decision was made to maximize the potential level of participation in these alternative survey modes and to minimize the potential household burden and avoid the confusion of self-administering a within-household randomization protocol.

2.5. Protocols for Telephone Contacting

Standard BRFSS data collection protocols specify that all unanswered telephone numbers and contacted households that did not refuse the survey receive 15 callbacks on different days of the week and times of the day or by appointment. This protocol was followed for the baseline data collection but the maximum number of calls was reduced to 10 in follow-up of nonrespondents. In both the baseline data collection and the web and mail surveys, attempts were made to convert initial refusals, excluding hostile refusals, to acceptance of the survey. All other BRFSS data collection protocols were in place for telephone follow-up of nonrespondents.

While CATI follow-up began on the 10th of the month, completed web and mail surveys were accepted through the end of the month. Additionally, if a respondent happened to complete both a self-administered questionnaire and a CATI interview, the self-administered responses were used in the analyses presented here.

3. Results

In total, 6,117 interviews were completed: 3,124 for the CATI-only data collection, 1,798 for the web experiment (1,160 web surveys and 638 CATI follow-ups), and 1,195 for the mail experiment (870 mail surveys and 325 CATI follow-ups). Analyses of these data included examination of differences in response rates across modes, the characteristics of study participants by experiment group and survey mode, and the reasons given by subjects for not participating in the web and mail surveys.

3.1. Response Rates

3.1.1. Address-matched Subjects

A central question addressed by this research is whether the use of alternative modes (web or mail) in conjunction with telephone (CATI) follow-up of nonrespondents can produce higher response rates than those obtained by using CATI only. To examine this question, we first compared the response rates in the web and mail mixed-mode surveys with the rates for the address-matched portion of the baseline study. Response rates were calculated by using the American Association for Public Opinion Research response rate formula #4 (AAPOR 2004).

In the address-matched sample, the response rates for the baseline data collection with CATI only ranged from 31.3 percent (New York) to 49.5 percent (North Dakota); the average rate was 40.1 percent across the four states (Table 1). Both the web survey with CATI follow-up and the mail survey with CATI follow-up had significantly higher

Table 1. Response rates for address-matched sample, by state and survey mode

State	Baseline	Web survey experiment			Mail survey experiment			Significance		
	CATI-only	Web-only +	CATI follow-up =	Overall web + CATI follow-up	Mail-only +	CATI follow-up =	Overall mail + CATI follow-up	(1)	(2)	(3)
State mean	40.1%	*15.4%*	*32.5%*	47.9%	*43.6%*	*16.4%*	60.0%	.001	.001	.001
(*n*)	(1,378)	*(1,905)*	*(1,905)*	(1,905)	*(501)*	*(501)*	(501)			
Arkansas	41.4%	*13.7%*	*34.0%*	47.7%	*37.8%*	*21.8%*	59.6%	.001	.001	.001
(*n*)	(1,314)	*(2,139)*	*(2,139)*	(2,139)	*(473)*	*(473)*	(473)			
Indiana	38.3%	*15.7%*	*32.3%*	48.0%	*43.3%*	*15.7%*	59.0%	.001	.001	.001
(*n*)	(1,518)	*(1,661)*	*(1,661)*	(1,661)	*(528)*	*(528)*	(528)			
New York	31.3%	*13.1%*	*24.7%*	37.8%	*39.7%*	*12.6%*	52.3%	.001	.001	.001
(*n*)	(1,833)	*(2,102)*	*(2,102)*	(2,102)	*(539)*	*(539)*	(539)			
North Dakota	49.5%	*19.2%*	*38.8%*	58.0%	*53.6%*	*15.3%*	68.9%	.001	.001	.001
(*n*)	(846)	*(1,720)*	*(1,720)*	(1,720)	*(464)*	*(464)*	(464)			

Note: Significance based on chi-square test with *p*-values noted for the following comparisons: (1) CATI baseline vs web survey with CATI follow-up, (2) CATI baseline vs mail survey with CATI follow-up, and (3) web survey with CATI follow-up vs mail survey with CATI follow-up.
Data for the web experiment CATI follow-up are weighted to adjust for subsampling of web nonrespondents.

response rates. Compared with response rates for CATI alone across the four states, rates for the web survey with follow-up averaged 8 percentage points higher and rates for the mail survey with follow-up averaged nearly 20 percentage points higher.

For the web and CATI mixed-mode design, 15.4 percent of respondents completed the web survey across the four states. An additional 32.5 percent were interviewed in the telephone follow-up. Use of the web survey was highest in North Dakota (19.2 percent) and lowest in New York (13.1 percent).

In the mail and CATI mixed-mode design, the percentage of completed interviews from the mail survey was larger than the percentage from the baseline telephone interviews in three of the four states (Indiana, New York, and North Dakota). On average, the response rate for the mail survey was 43.6 percent, compared with 40.1 percent for the CATI baseline. Response to the mail questionnaire was highest in North Dakota (53.6%) and lowest in Arkansas (37.8%). An additional 16.4 percent were interviewed in the CATI follow-up.

3.1.2. Overall Response Rates

The multimode approaches outperformed (in terms of response rates) the telephone-only approach in the address-matched sample. However, the more pressing question is whether use of multiple modes can produce higher rates of participation overall. A primary restriction on the use of web and mail as alternatives to telephone in an RDD survey is that the surveys are generally limited to households that can be reached by mail. Such surveys require household addresses, which can be obtained only by reverse-matching of telephone numbers to lists of known addresses or by making telephone contact with the household to get mailing information. Households for which an address cannot be obtained are not able, therefore, to participate in the web or mail surveys. These households do, however, need to be accounted for in overall response rate calculations.

We used data from the CATI baseline to estimate the effects on response rates of including non-address-matched households. Assuming that the final case disposition for the non-address-matched households would not have been significantly different if they had been contacted by telephone as part of the web or mail mixed-mode approaches, these baseline data can be used to calculate the expected overall response rates for the web and mail mixed-mode approaches. A weighting adjustment was used to ensure comparability across the groups by adjusting the proportion of address-matched households to match the proportion of address-matched households in the baseline sample.

Inclusion of non-address-matched households reduced differences in response rates among subjects selected for the baseline CATI, web, and mail surveys in each state (Table 2). In each state, the web and mail mixed-mode approaches produced higher response rates than CATI-only, but the impact of these approaches was tempered when compared to the rates for the address-matched-only cases. On average, web with CATI follow-up resulted in response rates that were 5 percentage points higher than the baseline CATI-only rates; mail with CATI follow-up resulted in response rates that were an average of 13 percentage points higher than baseline rates.

Table 2. Overall response rates, by state and survey mode

State	Response rate (%)		
	Baseline (CATI-only)	Web survey experiment	Mail survey experiment
State mean (*n*)	48.8 (1,673)	53.9 (1,686)	61.9 (1,734)
Arkansas (*n*)	50.4 (1,493)	53.9 (1,507)	62.7 (1,592)
Indiana (*n*)	47.1 (1,883)	53.6 (1,873)	60.4 (1,940)
New York (*n*)	39.2 (2,367)	43.4 (2,446)	52.4 (2,635)
North Dakota (*n*)	58.6 (950)	64.5 (920)	72.1 (967)

Note: statistical significance based on chi-square test. Differences in response rates across all columns in each row are significant at $p < .05$.
Estimated eligible households shown in parentheses (*n*).

3.2. Respondent Demographics

Do increases in response rates for the mixed-mode approaches lead to a different mix of respondents who complete interviews? More importantly, can the use of the alternative modes lead to increased participation among subgroups that are traditionally underrepresented in telephone studies (e.g., men, nonwhites, and younger individuals). We examined this question, comparing selected demographic characteristics of respondents across the baseline group and two experimental groups. For reference, we also compared population estimates for these groups, as determined by the U.S. Census Bureau's 2002 American Community Survey.

3.2.1. Comparison of CATI-only and Mixed-mode Approaches

The baseline CATI data support the supposition that in terms of unweighted demographic characteristics, BRFSS data – like those of many other RDD surveys of the general population – tend to overrepresent certain subgroups, particularly women, whites (non-Hispanics), and upper socioeconomic groups (Table 3). Compared with population-based estimates from the American Community Survey, higher percentages of respondents to the CATI-only mode were women, were white (non-Hispanic), had higher levels of education, and had higher family income than the adults in the general populations of these four states. Additionally, the subset of respondents from address-matched CATI households tended to overrepresent these groups to an even greater extent than the non-address-matched sample. The address-matched group also had a higher percentage of persons aged 65 years or older and a lower percentage of those aged 18 to 34 years than is seen in the general population.

Because respondents to the web survey and the mail survey were drawn only from address-matched households, it is not surprising that they overrepresented the same demographic groups that tend to be overrepresented in RDD surveys generally (Table 3). Moreover, for some characteristics, the percentage of particular subgroups among the web or mail survey (or both) respondents was significantly higher than the percentage of subjects in the baseline address-matched subgroup. For instance, respondents to the web survey with CATI follow-up were significantly more likely than baseline address-matched CATI respondents to be women, to be white (non-Hispanic), to have at least some college

Table 3. Demographic characteristics of respondents, by survey mode

Demographic characteristics	American Community Survey 2002 %	Baseline CATI-only (no address) %	Baseline CATI-only (with address) %	Web + CATI follow-up (with address) %	Mail + CATI follow-up (with address) %	Significance		
						(1)	(2)	(3)
Sex						.05	.01	.05
Male	48.8	35.2	41.9	38.9	35.7			
Female	51.2	64.8	58.1	61.1	64.3			
(*n*)		(1,008)	(2,146)	(3,611)	(1,195)			
Race						.05	n.s.	n.s.
White, non-Hispanic	78.9	82.3	87.5	90.6	89.7			
Other	21.1	17.7	12.5	9.4	10.3			
(*n*)		(1,004)	(2,130)	(3,591)	(1,189)			
Age						n.s.	.001	.01
18–34	30.3	29.8	19.6	17.8	13.3			
35–54	39.4	39.7	39.7	40.1	38.0			
55–64	13.0	14.7	17.2	16.0	18.0			
65+	17.3	15.9	23.6	26.1	30.7			
(*n*)		(1,001)	(2,121)	(3,605)	(1,182)			
Education						.001	n.s.	n.s.
High school or less	49.8	46.7	43.9	39.9	40.9			
Some college or more	50.2	53.3	56.1	60.1	59.1			
(*n*)		(1,008)	(2,141)	(3,603)	(1,173)			
Income						.05	n.s.	.05
< $50,000	61.7	48.3	45.0	42.1	47.7			
$50,000+	38.3	51.7	55.0	57.9	52.3			
(*n*)		(875)	(1,884)	(3,208)	(1,060)			
Have children						n.s.	.05	n.s.
None	69.2	60.2	67.0	68.9	70.5			
One or more	30.8	39.8	33.0	31.1	29.5			
(*n*)		(1,007)	(2,145)	(3,599)	(1,186)			

Note: Statistical significance based on chi-square test with *p*-values for the following comparisons: (1) CATI baseline vs web survey with CATI follow-up, (2) CATI baseline vs mail survey with CATI follow-up, and (3) web survey with CATI follow-up vs mail survey with CATI follow-up. Population estimates from the American Community Survey 2002 and sample demographics for non-address-matched subjects in the CATI baseline study are provided for reference.

education, or to have family income of at least $50,000 per year. Likewise, a higher percentage of respondents to the mail survey with CATI follow-up were women and/or adults aged 65 years or older than were those interviewed by CATI only.

Direct comparison of the two mixed-mode approaches showed significant differences across the two groups as well. Compared with respondents in the web experiment, the respondents in the mail experiment had a significantly higher percentage of women and adults aged 65 years or older. Compared with respondents in the mail experiment, the respondents in the web experiment were more likely to have incomes of at least $50,000 per year.

3.2.2. Comparison of Self-administered and Interviewer-administered Modes

Next, we compared respondents to the self-administered modes with those interviewed in the CATI follow-ups. The percentage of women responding to the self-administered surveys was significantly higher than the percentage responding to the respective CATI follow-ups: web survey = 63.4 percent, CATI follow-up = 60.0 percent; mail survey = 65.3 percent, CATI follow-up = 61.8 percent (Table 4). Respondents to the web survey and the mail survey were more likely to be white (non-Hispanic) than were those interviewed in each CATI follow-up: web survey = 93.0 percent, CATI follow-up = 89.5 percent; mail survey = 91.6 percent, CATI follow-up = 86.1 percent. Likewise, respondents to the self-administered modes were far more likely to have some college or more education than were those interviewed in the telephone follow-up. The percentage of web survey respondents having some or more college was greater than the corresponding percentage of mail survey respondents: web survey = 76.4 percent, CATI follow-up = 52.4 percent; mail survey = 62.4 percent, CATI follow-up = 50.6 percent.

There were also significant differences for several other characteristics across modes, but the pattern varied across the web and mail experiments. First, subjects responding to the web survey were significantly more likely to be of middle age (35 to 54 years) than were those in the CATI follow-up. Conversely, respondents to the mail surveys were much more likely to be aged 65 years or older and less likely to be ages 18 to 34 years than were those interviewed in the CATI follow-up. Second, the percentage of web survey respondents with income of at least $50,000 was significantly higher than the corresponding percentage in the CATI follow-up (web survey = 71.6 percent, CATI follow-up = 50.6 percent). There were no significant differences in the income of mail survey respondents and the respective CATI follow-up respondents. Third, the web survey respondents were significantly more likely than those in the CATI follow-up to live in households with one or more children, and respondents to the mail survey were more likely to live in households with no children, a finding that correlates with the tendency of these respondents to be older.

3.3. Reasons for not Participating in Self-administered Modes

Why did subjects choose not to complete the self-administered questionnaire during the first phase of data collection, and are there systematic patterns to these responses across population subgroups? Information was collected during the CATI follow-up for those

Table 4. Demographic characteristics of respondents, by experiment and survey mode

Demographic characteristics	Web survey with CATI follow-up		Mail survey with CATI follow-up		Significance (*p*-value)			
	Web	CATI	Mail	CATI	(1)	(2)	(3)	(4)
Sex					.141	.278	.400	.577
Male	36.6	40.0	34.7	38.2				
Female	63.4	60.0	65.3	61.8				
(*n*)	(1,159)	(638)	(870)	(325)				
Race					.01	.01	.235	.137
White, non-Hispanic	93.0	89.5	91.6	86.1				
Other	7.0	10.5	8.4	13.9				
(*n*)	(1,148)	(636)	(865)	(324)				
Age					.001	.001	.001	.079
18–34	15.3	18.8	10.3	21.1				
35–54	51.4	34.5	38.6	37.0				
55–64	18.5	14.9	17.9	18.0				
65+	14.8	31.7	33.2	23.9				
(*n*)	(1,157)	(637)	(861)	(322)				
Education					.001	.001	.001	.633
High school or less	23.6	47.6	37.6	49.4				
Some college or more	76.4	52.4	62.4	50.6				
(*n*)	(1,155)	(637)	(849)	(324)				
Income					.001	1.000	.001	.712
< $50,000	28.4	49.4	47.6	47.8				
$50,000 +	71.6	50.6	52.4	52.2				
(*n*)	(1,049)	(561)	(785)	(274)				
Marital status					.001	.001	.001	.891
Not married/single	24.8	45.1	33.6	45.7				
Married/couple	75.2	54.9	66.4	54.3				
(*n*)	(1,154)	(637)	(860)	(324)				
Have children					.001	.05	.001	.102
None	63.0	71.7	71.9	66.4				
One or more	37.0	28.3	28.1	33.6				
(*n*)	(1,154)	(636)	(862)	(324)				

Note: Statistical significance based on chi-square test with *p*-values for the following comparisons: (1) web survey vs web CATI follow-up, (2) mail survey vs mail CATI follow-up, (3) web survey vs mail survey, and (4) web CATI follow-up vs mail CATI follow-up.

who did not complete the web or mail survey. At the start of the CATI follow-up interview subjects were asked a short series of questions about the self-administered survey. These questions included (a) whether they remembered receiving the prenotification mailings; (b) why they chose not to complete the web or mail survey; and (c), for web survey sample members only, whether they have access to the Internet at home or elsewhere.

Of the 638 nonrespondents to the web survey, 37.5 percent said they did not respond because they did not remember receiving the mailings, nearly one-third (32.9 percent) indicated that they did not have access to the Internet, and almost 12 percent said they had no time to complete the survey (Table 5). Approximately 18 percent gave some other reason, such as not being comfortable using the web, security concerns, transmission speed, or not being interested in surveys. These responses varied significantly across subgroups. For example, lack of access to the Internet was the response given most often by women, those age 65 years or older, individuals with lower levels of education or family income below $50,000 per year, and those with no children in the household. Conversely, a higher percentage of men, respondents aged 18 to 34 years, those with some college or more or incomes of at least $50,000, and those with one or more children in the household said they did not complete the self-administered survey because they did not receive the mailing materials. Additionally, given the percentage of CATI completes obtained from web nonrespondents, it appears that few if any sample members who do not have access to the web were offended by the initial receipt of a web-only survey and refused to complete the survey for this reason.

Of the 324 mail survey nonrespondents who answered the CATI follow-up, a majority (57.7 percent) said they did not remember receiving the mailings, one in five (19.1 percent) said they had no time, and nearly one-quarter (23.1 percent) gave some other response. Some of the "other" responses included losing or misplacing the questionnaire, concern over the receipt of "junk mail," and a general lack of interest in participating in surveys. A majority of respondents in all of the subgroups examined, with the sole exception of those 65 years and older, said that not receiving the letters was the primary reason they did not complete the mail survey. Among those 65 and older, nearly 47 percent said they did not receive the letter and 14.3 percent said they had no time. However, a substantial proportion (39 percent) gave some other reason; "misplaced or lost the questionnaire" and "not interested in surveys" were the most frequently given responses.

4. Discussion

These mixed-mode experiments show that web surveys and mail surveys with telephone follow-up of nonrespondents are both possible alternatives to the current CATI-only approach for significantly increasing participation in an RDD sample survey of the adult population. Both web and mail surveys in conjunction with CATI follow-up of nonrespondents produced higher response rates than the CATI-only approach. Moreover, these findings held in each of the four states examined. The mixed-mode approaches, mail in particular, generated roughly equivalent percentage point increases across the four states, even in North Dakota, which had by far the highest telephone-only response rate. This finding suggests that there is considerable room for improvement in response rates, regardless of whether state residents are typically more likely to participate.

Table 5. Reasons for CATI respondents not completing Web/mail self-administered surveys, by demographic characteristics

Demographic characteristics	Web survey experiment						Mail survey experiment				
	(*n*)	No access to Web (%)	Letter not received (%)	No time (%)	Other reason (%)	Sig.	(*n*)	Letter not received (%)	No time (%)	Other reason (%)	Sig.
Total	638	32.9	37.5	11.9	17.7		324	57.7	19.1	23.1	
Sex						.001					.02
Male	256	28.9	46.9	9.0	15.2		124	66.1	18.5	15.3	
Female	382	35.6	31.2	13.9	19.4		200	52.5	19.5	28.0	
Race						n.s.					n.s.
White, non-Hispanic	569	33.0	36.9	12.1	17.9		278	55.4	20.1	24.5	
Other	67	31.3	43.3	10.4	14.9		45	71.1	13.3	15.6	
Age						.001					.01
18–34	120	5.0	65.8	13.3	15.8		68	73.5	14.7	11.8	
35–54	220	20.5	42.7	17.7	19.1		119	56.3	23.5	20.2	
55–64	95	26.3	30.5	15.8	27.4		57	54.4	22.8	22.8	
65+	202	65.8	18.3	3.0	12.9		77	46.8	14.3	39.0	
Education						.001					n.s.
High school or less	303	49.5	29.0	6.3	15.2		160	57.5	16.9	25.6	
Some college or more	334	17.7	45.2	17.1	20.1		163	57.7	21.5	20.9	
Income						.001					.01
< $50,000	277	47.7	28.9	5.4	18.1		130	54.6	14.6	30.8	
$50,000+	284	14.8	48.2	19.7	17.3		143	62.9	23.1	14.0	
Have children						.001					n.s.
None	456	41.4	29.2	9.2	20.2		215	54.4	18.6	27.0	
One or more	180	11.7	57.8	18.9	11.7		108	64.8	19.4	15.7	

Note: Statistical significance based on chi-square test.

The modest increase in the overall response rate for the web mixed-mode approach needs to be viewed, however, with caution. Recent experiments in the use of advance letters with similar populations showed that prenotification alone can increase response rates by approximately 6 percent (Link and Mokdad 2005a). Thus, the 5-percent increase in response rates obtained in the web experiments may have been due more to use of mail as a means of notifying potential web respondents about the study as opposed to the use of an alternative mode to CATI. Additionally, the completed web surveys made up a much smaller percentage of the total surveys completed in the web-CATI mixed-mode experiment, compared with the percentage of mail surveys completed in the second experiment. Web survey respondents accounted for one-third of the completed interviews in that experiment (15.4 percent of 47.9 percent total); mail survey respondents accounted for nearly three-quarters of the completed interviews in that experiment (43.6 percent of 60 percent total).

In contrast to use of the web, use of a self-administered hardcopy survey resulted in substantial improvement in response rates. Overall, the mail experiment produced an improvement of 13 percentage points over the baseline CATI data collection, far exceeding the 6-percent gain reported for the use of advance letters alone. The mail questionnaire, like the web survey, did not, however, contain the entire set of BRFSS questions typically asked in most of the states. The experiments were confined to the core component and did not include any of the optional question modules or state-specific modules. It is likely that extending the length of the mail version of the questionnaire could have a negative effect on the resulting response rates. Additional research is necessary, therefore, to determine a productive cutoff for questionnaire length. How long is too long, making the length of the mail survey detrimental in light of the obtainable response rates?

Although both mixed-mode approaches increased response rates, the increases were not even across demographic groups. In fact, the designs tested here actually exacerbated the problems associated with overrepresentation of particular subgroups of the population. For example, telephone surveys typically overrepresent women who have completed surveys, because men are in general more difficult to contact at home and tend to be more reluctant to participate in surveys. One common hypothesis is that more men might be encouraged to participate in surveys like the BRFSS if given the option of responding to a self-administered survey at a convenient time of their choosing. This was not the case. In this study, the percentage of women completing the web survey or the mail survey was significantly higher than the percentage of men completing either survey. Likewise, use of a web survey was initially hypothesized to encourage greater participation by younger respondents, especially those ages 18 to 34 years. Again, this was not the case. Web respondents were more likely to be middle-aged (35 to 54 years) than were those completing either the mail survey or CATI alone. In contrast, compared with respondents to other survey modes, a higher percentage of mail survey respondents were 65 years and older.

Certainly data can be adjusted so that the sample demographics of data collected by these different modes better reflect the characteristics of the general population. However, weighting and post-stratification are often limited to a few key demographic variables for which population estimates are available. Any differences between respondents and nonrespondents that are not corrected for through these procedures can lead to potential

bias in the estimates obtained from the survey. Another less desirable by-product of weighting is that when the variance of the weights is large, the resulting standard errors are often also larger than they would be for unweighted estimates. As a result, the precision of point and parameter estimates will often decline (Groves et al. 2004).

Why did the use of these survey modes increase the gap in participation among certain subgroups? As with any research, study design has an important effect in determining study results. In this set of experiments, two design features appear to have influenced the outcomes, particularly in relation to the types of individuals who responded using the self-administered modes. First, findings on the use of reverse matching to databases to identify addresses for subjects show that the demographic characteristics of the address-matched sample tended to be less representative of the adult general population than the characteristics of the non-address-matched sample (Link and Mokdad 2005a). Drawing the web and mail survey samples only from the address-matched households increased the likelihood, therefore, that these subgroups would be overrepresented by these modes as well. Future research should explore ways to offer alternative survey modes to subjects in the non-address-matched group. Collecting address information or offering information on web access over the telephone is one way to reach this group. As an alternative to mailing, households could be contacted by telephone and asked to participate. Then they could be (1) provided with a user name and password for the web by telephone or mail, or (2) sent a questionnaire if an address was provided. This approach has the advantage of allowing contact with both address-matched and non-address-matched households. One significant obstacle, however, is the difficulty of obtaining cooperation from households in RDD-based telephone studies, let alone cooperation enough to explain the study, elicit participation, and obtain mailing addresses for the purposes of sending additional study information to the sample member. Additionally, if cost saving is a goal, assuming the self-administered modes cost less than telephone interviewing, this approach may not be feasible. Likewise, if time in the field is an issue, as it is with BRFSS which requires monthly collection of data, then time pressure may make this approach less effective.

An alternative may be to move away from reliance on RDD sampling frame for selecting potential responders for alternative modes. A number of commercial vendors offer web-sampling frames, however, to date none are able to offer reliable probability-based frames. Likewise, questions about the completeness and representativeness of most mail survey frames have limited the widespread use of mail surveys for general population surveys. Recent studies have shown, however, that some lists (e.g., the Delivery Sequence File offered by the U.S. Postal Service) may provide coverage similar to that of telephone-based approaches (Iannacchione, Staab, and Redden 2003).

The second design feature that likely affected who participated was the method of within-household selection. By allowing households completing the web and mail surveys to self-select a participant, we likely reduced the potential confusion and frustration associated with self-administering a random-selection technique and increased participation in both mixed-mode experiments. However, this approach probably contributed to the substantial increase in participation among women and older individuals. Historically in telephone surveys, women often volunteer to complete a survey even when a man in the household is randomly selected to participate. This is likely here as well and may have contributed also to the finding that men were more likely in the

telephone interviews to report never having seen the web or mail survey materials mailed to the home. A similar situation may exist with younger adults, particularly those living at home with older parents. A package addressed to the household is more likely to be responded to by the parents, especially the mother, than by the younger adult. Additionally, adults aged 18 to 34 years are also more mobile, changing addresses more readily than older individuals. As a result, the increased likelihood of adults in this age group reporting that they did not see the mailed materials could also be attributable to the subject changing his or her address but keeping the same telephone number. Future studies should test ways to balance the trade-off between reduced nonresponse and the need for techniques for within-household selection of subjects that can produce a more representative group of survey respondents.

Finally, there remains the issue that some subgroups of the population choose not to participate in surveys at the same rate as other subgroups. Men, younger adults, and nonwhites have historically made this choice. Unfortunately, the use of mixed-mode approaches in our study did not help to address this issue. Lack of access to the Internet is a constraint in the use of that mode, particularly in efforts to reach nonwhites and persons of low socioeconomic status. In addition to the suggestions outlined here, greater attention needs to be given to determining the types of appeals, modes, and approaches that work most effectively in motivating these underrepresented groups to participate in surveys.

In conclusion, the results of these experiments show that the use of web surveys and mail surveys in conjunction with more traditional CATI may be viable alternatives for increasing participation rates in the BRFSS, but not without the potential for exacerbating the differences between respondents and nonrespondents and possibly increasing nonresponse bias in the estimates derived from these data. Considerable work will be required, therefore, before a sound recommendation can be made to use a multiple-mode approach in an on-going, telephone-based surveillance like the BRFSS. Perhaps most important is the need to recognize that issues involving the use of new methods and technologies (e.g., the Internet) must continue to be studied. Technology and communications are changing at such a rapid rate that the research findings of today may not be relevant for long. Survey methods and health surveillance techniques are in a constant state of evolution, and it is incumbent on researchers and methodologists to stay abreast of these changes, embracing through a rigorous process of testing and validation the new technologies and approaches that can improve the quality of the data and estimates we produce.

5. References

Aquilino, W.S. (1994). Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment. Public Opinion Quarterly, 58, 210–240.

American Association for Public Opinion Research (2004). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Ann Arbor, MI: AAPOR.

Babbie, E. (1990). Survey Research Methods. Belmont, CA: Wadsworth Publishing Company.

Benson, G., Booman, W.P., and Clark, K.E. (1951). A Study of Interview Refusals. Journal of Applied Psychology, 35, 116–119.

Bryant, B. (1975). Respondent Selection in a Time of Changing Household Composition. Journal of Marketing Research, 12, 129–135.

Church, A.H. (1993). Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-analysis. Public Opinion Quarterly, 57, 62–79.

Converse, J.M. (1976). Predicting No Opinion in the Polls. Public Opinion Quarterly, 40, 515–530.

Coombs, C.H. and Coombs, L.C. (1976). Don't Know: Item Ambiguity or Response Uncertainty? Public Opinion Quarterly, 40, 497–514.

Couper, M.P. (2000). Web Surveys: A Review of Issues and Approaches. Public Opinion Quarterly, 64, 464–494.

De Leeuw, E.D. (1992). Data Quality in Mail, Telephone, and Face-to-Face Surveys. Amsterdam: TT Publications.

De Leeuw, E.D. and De Heer, W. (2002). Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In Survey Nonresponse, R.M. Groves, D.A. Dillman, J.L. Eltinge, and R. Little (eds). New York: John Wiley and Sons.

DeMaio, T.J. (1984). Social Desirability and Survey Measurement: A Review. In Surveying Subjective Phenomena: Volume 2, C.F. Turner and E. Martin (eds). New York: Russell Sage.

Dillman, D.A. (1999). Mail and Other Self-administered Surveys in the 21st Century: The Beginning of a New Era. The Gallup Research Journal, Winter/Spring, 121–140.

Dillman, D.A. (2000). Mail and Internet Surveys: The Tailored Design Method. New York: John Wiley and Sons, Inc.

Dillman, D.A., Eltinge, J.L., Groves, R.M., and Little, R. (2002). Survey Nonresponse in Design, Data Collection, and Analysis. In Survey Nonresponse, R.M. Groves, D.A. Dillman, J.L. Eltinge, and R. Little (eds). New York: John Wiley and Sons.

Dillman, D.A., Sangster, R.L., Tanari, J., and Rockwood, T. (1996). Understanding Differences in People's Answers to Telephone and Mail Surveys. In New Directions for Evaluation Series, 70, Advances in Survey Research, M.T. Braverman and J.K. Slater (eds). San Francisco: Jossey-Bass.

Dillman, D.A., Gallegos, J., and Frey, J. (1976). Reducing Refusal Rates for Telephone Interviews. Public Opinion Quarterly, 40, 66–78.

Ford, N.M. (1967). The Advance Letter in Mail Surveys. Journal of Marketing Research, 9, 202–204.

Faulkenberry, G.D. and Mason, R. (1978). Characteristics of Nonopinion and No Opinion Response Groups. Public Opinion Quarterly, 42, 533–543.

Goldstein, K. and Jennings, M. (2002). The Effect of Advance Letters on Cooperation in a List Sample Telephone Survey. Public Opinion Quarterly, 66, 608–617.

Groves, R. (2002). Principles and Practices in Non-response Reduction. Presentation at the 2002 Respondent Cooperation Workshop sponsored by the Council for Marketing and Opinion Research, New York, NY.

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., and Tourangeau, R. (2004). Survey Methodology. New York, NY: Wiley.

Groves, R.M. and Kahn, R.L. (1979). Surveys by Telephone: A National Comparison with Personal Interviews. New York: Academic.

Iannacchione, V.G., Staab, J.M., and Redden, D.T. (2003). Evaluating the Use of Residential Mailing Addresses in a Metropolitan Household Survey. Public Opinion Quarterly, 67, 202–210.

James, J.M. and Bolstein, R. (1990). The Effect of Monetary Incentives and Follow-up Mailings on the Response Rate and Response Quality in Mail Surveys. Public Opinion Quarterly, 54, 346–361.

James, J.M. and Bolstein, R. (1992). Large Monetary Incentives and Their Effect on Mail Survey Response. Public Opinion Quarterly, 56, 442–452.

Kennedy, J.M., Parks, R.B., Bannister, N., Inghram, J., and Terhune, H. (1998). An Analysis of the Effect of Varying Presurvey Letter Characteristics on Cooperation Rates. Paper presented at the annual conference of the American Association for Public Opinion Research, St. Louis, MO.

Link, M.W. (2003). BRFSS Response Rates: Which Way is UP? Paper presented at the Behavioral Risk Factor Surveillance System Annual Conference, St. Louis, MO.

Link, M.W. and Mokdad, A.H. (2005a). Use of Prenotification Letters: An Assessment of Benefits, Costs, and Data Quality. Public Opinion Quarterly, 69, 572–587.

Link, M.W. and Mokdad, A.H. (2005b). Effects of Survey Mode on Self-Reports of Adult Alcohol Consumption: Comparison of Web, Mail, and Telephone Approaches. Journal of Studies on Alcohol, 66, 239–245.

Link, M.W. and Mokdad, A.H. (2005c). Use of Alternative Modes for Health Surveillance Surveys: Results from a Web/Mail/Telephone Experiment. Epidemiology, 16, 701–704.

Link, M.W., Armsby, P., Hubal, R. and Guinn, C. (2006). Accessibility and Acceptance of Responsive Virtual Human Technology as a Telephone Interviewer Training Tool. Computers in Human Behavior, 22, 412–426.

McCabe, S.E., Boyd, C.J., Couper, M.P., Crawford, S., and D'Arcy, H. (2002). Mode Effects for Collecting Alcohol and Other Drug Use Data: Web and U.S. Mail. Journal of Studies on Alcohol, 63, 755–761.

Schonlau, M., Asch, B.J., and Du, C. (2003). Web Surveys as Part of a Mixed Mode Strategy for Populations That Cannot Be Contacted by E-mail. Social Science Computer Review, 21, 218–222.

Singer, E., Hoewyk, J., and Maher, M. (2000). Experiments with Incentives in Telephone Surveys. Public Opinion Quarterly, 64, 171–188.

Singer, E., Groves, R.M., and Corning, A. (1999). Differential Incentives: Beliefs and Practices, Perceptions of Equality, and Effects on Survey Participation. Public Opinion Quarterly, 63, 251–260.

Received March 2004
Revised February 2005