

**THE NET EFFECT:
A COMPARISON OF INTERNET AND MAIL SURVEY RESPONDENTS**

Esther M. Friedman, Nancy A. Clusen, Mathematica Policy Research
Michael Hartzell, Department of Defense
Esther M. Friedman, 1606 Calibre Woods Drive NE Atlanta, GA 30329 (emf28@columbia.edu)

Key Words: Web Survey; Internet; Health Care Survey of Department of Defense Beneficiaries

1. Introduction

Multiple-mode surveys can have the potential to improve response and reduce survey error and costs (Groves, 1989). Because of the relatively low costs of implementation and the increased prevalence of internet access in the general population, Web surveys are an increasingly popular mode of survey administration. However, as some individuals do not yet have internet access and levels of computer literacy vary, web surveys are often combined with other surveys, particularly mail surveys, to ensure that important subsets of the population are not excluded. The full effect Web data collection may have on a primarily mail survey is not yet fully understood. Some characteristics of a Web survey are similar to those of a mail survey (self-administration and presenting the questions visually), while others are unique to a survey implemented via the Web (accessing the instrument with a username and password and viewing only a portion of the questionnaire at one time). Web instruments also allow for formats and designs that are not possible in other modes, and the impacts of these formatting decisions on response error (Couper, et al., 2004) and potential mode of administration effects (Manfred and Vehovar, 2002; Chatt and Dennis, 2004) are only now beginning to be explored. Moreover, Web survey design and Web instrument formatting may affect unit and item nonresponse and data quality (Couper, 2000; Couper et. al., 2001; Porter and Whitcomb 2003).

There are many contradictory findings in the survey literature regarding web respondents. In a survey of university students, Kwak and Radler (2002) found a Web survey to have a lower response rate than a mail survey, as well as lower item nonresponse and longer open-ended responses. They also found demographic differences between the two modes. However, in a survey of Census Bureau employees, Bates (2001) did not find significant differences between the demographic characteristics of mail and internet respondents to. She also found that Internet respondents have a higher rate of item nonresponse than do mail respondents. Another study of university students showed that Web surveys can achieve a comparable response rate to a mail survey (Kaplowitz et. al., 2004).

However, age differences between the two modes remain. Moreover, though one reason for adding a web option is to improve response, Griffin et. al. (2001) found that it does not necessarily improve response and that it actually significantly lowered the response rates. They hypothesized that adding an internet mode of response to a mail survey contributed to a break in the response process.

This paper details response rates by Web and mail in a mail survey with a web response option, and compares the characteristics of beneficiaries who chose to respond online to those of beneficiaries who responded by mail. The purpose is to determine whether Web respondents differ from mail respondents and whether particular subsets of the population are more likely to respond by Internet.

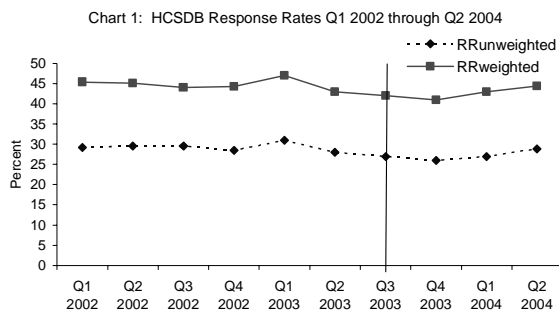
2. Method

The Adult Health Care Survey of Department of Defense Beneficiaries (HCSDB) is designed to monitor experience with, access to, and use of care provided by the Military Health System (MHS). The HCSDB has been conducted annually since 1995, and was first fielded quarterly in 2001. The samples for the HCSDB quarterly surveys were selected using a stratified sample design, stratified by geography, enrollment status, and beneficiary group. In order to account for the survey design and to control for the biasing effects of nonresponse, weights were calculated for each quarter of data. The final weight includes a sampling weight that reflects the differential selection probabilities used to sample beneficiaries across strata and weighting class adjustments for nonresponse. The quarterly surveys were mailed to representative, independently selected samples of 45,000 MHS beneficiaries per quarter in 2003 and 50,000 MHS beneficiaries in 2004. The mailing consists of a notification letter (mailed about a month before the survey), questionnaire, reminder/thank you postcard (mailed two weeks after the first mailing), and a follow-up questionnaire, with the entire fielding period lasting eight weeks. Beginning in quarter three of 2003, beneficiaries were given the option of responding Web in both the first and second survey cover letter. The cover letters included a link to a Website and a unique password.

3. Response Rates

The Web survey was introduced primarily to increase response, particularly for active duty beneficiaries, a group with a historically low response rate and easy access to the internet. Some studies have indicated that introducing an additional mode of response may not only fail to increase response, but may even decrease it (Griffin, Fischer, and Morgan, 2001). A possible explanation for this is that potential respondents may intend to respond using one of the modes, put off deciding which one, and then may forget to respond at all. This is certainly likely in the case of the HCSDB, where beneficiaries who decide to respond by Web must save the link and password and take the additional step of logging on to a computer at a later time and filling out the questionnaire. It is certainly possible for such beneficiaries to lose the log-in information, or to have difficulty logging on to the website and give up filling out the Web survey after having already discarded the paper version or having forgotten that there was a paper version at all. Furthermore, Keesling (2002) studied several techniques for improving web response and awareness in multi-mode surveys and found that actively advertising the availability of a web option helped improve web response over simply mentioning it in the cover letter. This is possibly because web surveys are not yet that common and individuals did not respond by web initially because they were not aware of the option. In the HCSDB survey, the web option was not specially advertised or underscored.

As chart 1 indicates, the response rates for the quarters with the introduction of the Web survey appear to be very similar to those of prior quarters over the past year. It is difficult, however, to determine whether these Web respondents are would-be nonrespondents, who are only responding now that the Web option is available, or if they are simply respondents who prefer the Web but who would respond nevertheless by any mode.



The percentage of beneficiaries who choose to respond by Web has been steadily increasing from quarter to quarter, with only 6 percent of respondents to the HCSDB choosing to respond by Internet in quarter

three of 2003, nine percent in quarter four, 10 percent in quarter one of 2004 and 13 percent in quarter two of 2004 (please see table 1 for further details). This may indicate that the HCSDB Web site is getting faster or, perhaps, that beneficiaries themselves are more likely to have faster Web access or to have Web access altogether. Another possible explanation is that people in general are becoming more comfortable with the internet and more accustomed to the idea of Web surveys and may have even filled out other Internet surveys already.

Table 1: Unweighted Response Rates Web Versus Mail

	Mail Respondents		Web Respondents	
	Count	Percent	Count	Percent
Q3, 2003	11,262	94%	753	6%
Q4, 2003	10,622	91%	1,010	9%
Q1, 2004	11,950	90%	1,332	10%
Q2, 2004	12,692	87%	1,807	13%

To explore this further, the percent of beneficiaries who logged on to the website but did not complete the questionnaire was calculated. As indicated in Table 2, 17 percent of beneficiaries who logged on to the Web survey in quarter 3, 2003 did not complete the questionnaire. This percentage dropped drastically to only four percent in quarter four, which seems to indicate that the problem may have been with the Web instrument itself being slow or difficult to use and that this was corrected in the next quarter. Typically, only one percent of mail questionnaires are partially completed, yet this discrepancy is not surprising as most people will not send back a partially completed mail questionnaire unless they believe it to be complete enough for use. It is plausible that more people will start to fill out a Web instrument but not finish it.

Table 2: Completes Versus Partial Completes, Web and Mail

	Potential Web Respondents		Potential Mail Respondents	
	Completes	Partial Completes	Completes	Partial Completes
Q3, 2003	83%	17%	99%	1%
Q4, 2003	96%	4%	99%	1%
Q1, 2004	93%	7%	99%	1%
Q2, 2004	96%	4%	99%	1%

4. Results

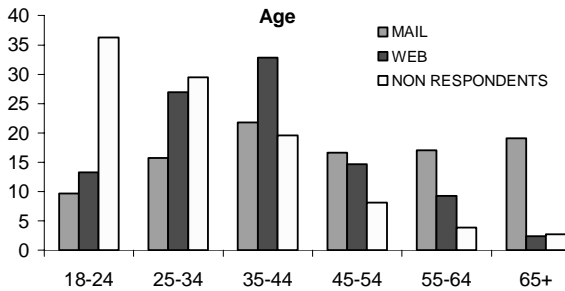
Many questions arise with the introduction of an internet response option, primarily including: Who chooses to respond by Web in place of mail? Perhaps these individuals are demographically different, or have different health-related characteristics from mail respondents, which both make them more likely to respond by Web and more likely to differ from mail respondents in terms of health indicators in the survey. Would these respondents have been nonrespondents had there been no Web option or would they have responded by mail? And, finally, since it is simple to

program skip logics and range checks into a web survey, do Web respondents exhibit better data quality than do mail respondents?

4.1 Demographic Variables

Data were analyzed for quarters three and four of 2003 and quarter one of 2004 to determine if there are any demographic differences between the two groups. Chart 2, below, depicts the percent of mail, Web, and nonrespondent for six age categories for quarter one of 2004 (similar results were found for quarters three and four of 2003). As the chart indicates, nonrespondents are most likely to be in the younger age categories with over 35 percent of beneficiaries sampled in quarter one of 2004, between the ages of 18 and 24 and only three percent of them over age 65. Interestingly, Web respondents are also more likely than mail respondents to be in these younger age categories. This is not surprising given that younger people are more likely to try out new technology and tend to be more computer-literate than their older counterparts. As Web respondents are more similar to nonrespondents than are mail respondents, these findings may also indicate that Web respondents might have been nonrespondents had a Web option not been available.

Chart 2: Percent Mail, Web, and Nonrespondent, by Age, Quarter 1 2004



Similarly, charts 3 through 7 treat several other demographic variables, namely, gender, region, sponsor's race (this is not necessarily the beneficiary's race, though it should be a good indicator), beneficiary group and education (education is a survey variable and information for nonrespondents is therefore unavailable). For all these variables, Web respondents differ from mail respondents and are more similar to nonrespondents than are mail respondents. As the charts below indicate, Web respondents are more likely than mail respondents to be male, to reside outside of the US, to be white, to be active duty beneficiaries, and to have at least some college, and are less likely to be retired.

Chart 3: Percent Mail, Web, and Nonrespondent, by Gender Quarter 1 2004

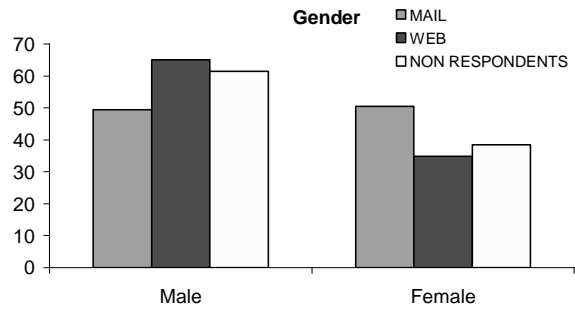


Chart 4: Percent Mail, Web, and Nonrespondent, by Region Quarter 1 2004

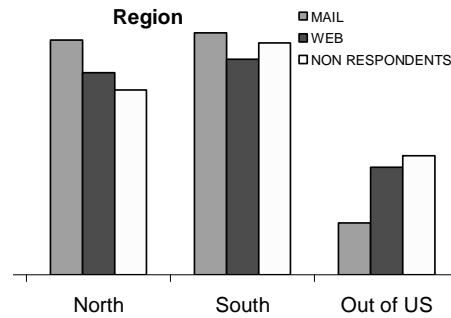


Chart 5: Percent Mail, Web, and Nonrespondent, by Sponsor's Race Quarter 1 2004

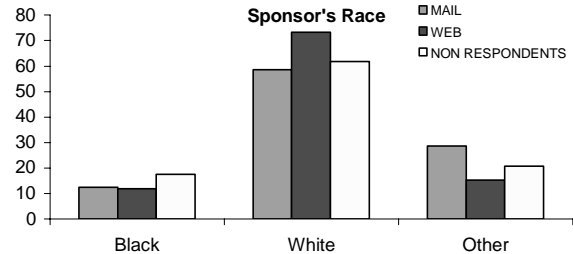


Chart 6: Percent Mail, Web, and Nonrespondent, by Beneficiary Group Quarter 1 2004

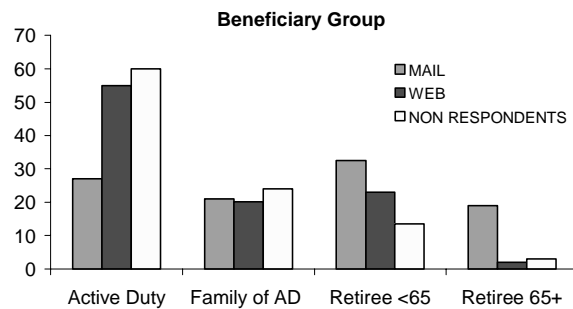
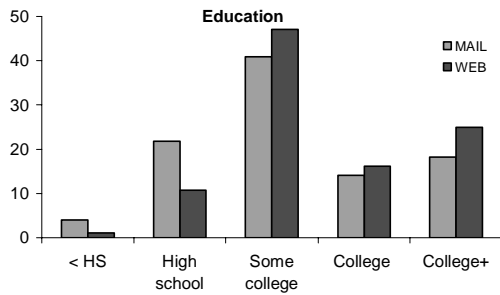


Chart 7: Percent Mail and Web, by Education, Quarter 1 2004



4.2 Self-Reported Health

As the primary purpose of the HSCDB is to measure beneficiary health usage, it is important to determine whether there are any health differences between the two groups. There are several reasons to hypothesize that mail survey respondents are less healthy than Web respondents. First, it is possible that less healthy individuals are not as mobile and may have difficulty getting to a computer to respond to an internet survey. Moreover, less healthy individuals may not have full-time employment and may therefore not have internet access at the workplace and may thus have limited internet availability. To test this hypothesis, several questions measuring the number of times in the last 30 days an individual felt that his physical or mental health generally was not good, felt that activity was hindered, or felt pain, felt sad/ depressed, etc., were analyzed using a logistic regression model, modeling Web response (vs. mail response). Since many demographic variables, such as age, clearly impact health and were found to be different for the two groups, the demographic variables discussed in section 4.1 were controlled for. However, as the results in Table 3 indicate, once all demographic variables were controlled for, there are no differences in self-reported health for the two groups for quarter one of 2004. There were, however, significant differences for the first two items in quarter three of 2003, with negative betas indicating that mail respondents may in fact have poorer health than Web respondents, as hypothesized. However, as this was not found to be the case for the other scales and as this was not found to be true for quarter 1, 2004 and quarter 3, 2003, it is difficult to know whether this is more than an anomaly.

Table 3: Beta Coefficients for Multivariate Logistic Regression for Self-Reported Health

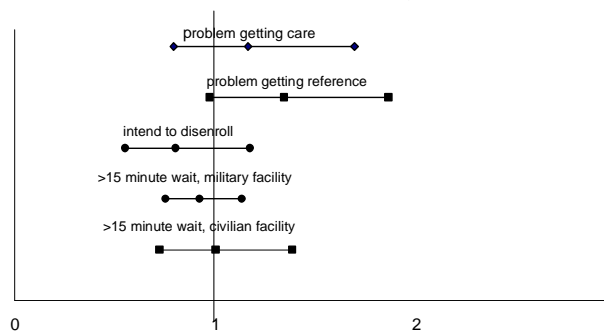
VARIABLE	Beta Coefficient		
	Q3, 2003	Q4, 2003	Q1, 2004
Last 30 days:Days physical hlth not good	-0.014 *	-0.004	0.001
Last 30 days:Days mental hlth not good	-0.013 *	-0.005	-0.002
Last 30 days:Days health hinder actvty	-0.005	-0.004	0.009
Last 30 days:Days pain hndr usual actvty	-0.009	-0.003	-0.001
Last 30 days:Days felt sad/depressed	-0.006	-0.011	-0.013
Last 30 days:Days wrried/tense/anxious	-0.005	-0.012	0.002
Last 30 days:Days without enough sleep	-0.005	0.001	-0.003
Last 30 days:Days hithy, full of enrgy	-0.002	0.002	0.004

* indicates significance at 0.05 level or better

4.3 Satisfaction with TRICARE

One reason for responding to a survey by mail may be dissatisfaction with the service being evaluated by the survey, in this case, TRICARE health insurance. Mail respondents, perhaps want to respond my mail so they have room to add additional comments and complaints. To test this hypothesis, several dichotomous variables measuring satisfaction (or lack thereof) with TRICARE are modeled using a logistic regression model. Once again, the demographic variables discussed above were controlled for in this model. However, as the results in Chart 8 indicate, there are no significant differences in satisfaction with TRICARE for quarter 1, 2004 or for the other two quarters tested: quarters three and four of 2003.

Chart 8: Odds ratios and 95% Confidence Intervals for web to mail respondents, Satisfaction with TRICARE Quarter 1, 2004



4.4 Quality of Response

There are several measures of data quality in the HSCDB survey. Respondents who return surveys with too many missing or incomplete values or with key items left blank will typically be treated as nonrespondents. However, data quality varies from survey to survey, even for respondents. One advantage of the Web survey is that checks may be built in to control quality of data by ensuring logical skip patterns and appropriate range checks. Moreover, whereas someone may accidentally skip a question in a mail survey or incorrectly fill in skip patterns, Web surveys

can remind people if things are missing and can prevent respondents from filling in questions that do not relate to their specific situation. Several variables measuring the number of responses of “don’t know,” and “not applicable” and the number of missing values were tested for significant differences using a logistic regression model, controlling for significant demographic variables. Not only was data quality found to be better for Web respondents for several variables in each quarter, but the number of significant variables increases from quarter to quarter. This is due to one of the primary advantages of a Web survey instrument: it is easy to add additional checks to each quarter and ensure reliable and accurate data. In quarter 1, 2004, blank responses were not even allowed for any survey questions. Caution should be used, however, in forcing responses, as respondents may give invalid responses solely to move on to the next question. In a study of Census Bureau Employees, Bates (2001) found that Internet respondents had a higher average percent of item nonresponse than do mail survey respondents. This may indicate that Web respondents to the HCSDB may have left more items blank if they had not been forced to respond. One should be cautious in interpreting results that may have been given in haste or frustration at being forced to respond to a given question.

Table 4: Beta Coefficients, Standard Errors, and Significance Indicators, Quality of Data

VARIABLE	Q3 2003	Q4 2003	Q1 2004
# response of DK	0.05	-0.13 *	-0.16 *
# responses n/a	-0.04 *	-0.08 *	-0.03 *
# blank responses	-0.01	-0.01	No blanks for Web!
total # missing responses(any reason)	-0.01 *	-0.02 *	-1.32 *

* indicates significance at 0.05 level or better

5. Conclusions

Adding a web option has not significantly affected response rates to the HCSDB, though with time, as people become more acquainted with internet surveys, this additional mode may help to increase response. Web respondents do, however, appear to differ from mail respondents and are more likely than mail respondents to be: young, White, male, active duty beneficiaries, and to live outside of the U.S. Their characteristics appear to be similar to those of nonrespondents, which may indicate that had there not been a Web option offered, some of them may not have responded at all. Furthermore, adding the web option may even help to decrease nonresponse bias by attracting a subset of the population that is more similar to nonrespondents to the study. For the most part, no differences were found in key estimates measuring health usage, self reported health, and satisfaction with TRICARE. Significant differences were, however, found in terms of data quality, with web respondents

having fewer blank responses or responses of “don’t know” or “not applicable.” This difference may, however, be due to additional checks and forced responses; further research should investigate whether these additional checks increase percent of response but decrease the validity of an item. Further research should also investigate whether advertising the web option will help to improve both web and overall response.

REFERENCES

Bates, N. (2001). “Internet versus Mail as a Data Collection Methodology from a High Coverage Population,” *Proceedings of the American Statistical Association Survey Research Methods Section [CD-ROM]*. Alexandria, VA: American Statistical Association.

Chatt, C. and Dennis, M. (2004). “Data Collection Mode Effects Controlling for Sample Origins in a Panel Survey: Telephone Versus Internet,” Paper presented at the Joint Statistical Meetings, Toronto, Canada.

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

Couper, M.P., Tourangeau R., and Kenyon K. (2004). “Picture This! Exploring Visual Effects in Web Surveys,” *Public Opinion Quarterly* 68:255-266.

Couper, M. P., Traugott M. W., and Lamias M. J. (2001). “Web Survey Design,” *Public Opinion Quarterly* 65:230-253.

Griffin, D.H., Fischer, D.P., and Morgan, M.T. (2001). “Testing an Internet Response Option for the American Community Survey,” Paper presented at the American Association for Public Opinion Research Conference, Montreal, Canada.

Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York: John Wiley & Sons.

Kwak, N. and Radler B. (2002). “A Comparison Between Mail and Web Surveys: Response Pattern, Respondent Profile, and Data Quality,” *Journal of Official Statistics* 18:257-273.

Keesling, R. (2002). “Use of Web-Based Data Collection to Supplement and Improve Mail Survey Response Rates,” Paper Presented at the American Association for Public Opinion Research Conference, St. Pete Beach, Florida.

Manfreda, K. L. and Vehovar V. (2002). "Mode Effect in Web Surveys," Paper Presented at the American Association for Public Opinion Research Conference, St. Pete Beach, Florida.

Porter, S. R. and Whitcomb M. E. (2003). "The Impact of Contact Type on Web Survey Response Rates," *Public Opinion Quarterly* 67:579-588.