Data Collection Mode Effects On Political Knowledge

Survey Methods: Insights from the Field

Mingnan Liu, Ph.D, Program in Survey Methodology, University of Michigan
Yichen Wang, NERA Economic Consulting

How to cite this article: Liu M. & Wang Y. (2014). Data collection mode effects on political knowledge, Survey Methods: Insights from the Field. Retrieved from https://surveyinsights.org/?p=5317

DOI: 10.13094/SMIF-2014-00009

Copyright: © the authors 2014. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

Abstract: Given the popularity of political knowledge questions in political science research, it is critical to understand the measurement of this question type. This study examines the data collection mode effect on political knowledge questions. Specifically, responses to political knowledge questions between face-to-face and Web surveys, and between computer-assisted self-interview (CASI) and Web surveys are compared using the 2012 American National Election Studies. The results suggest that a significant mode effect exists for political knowledge questions. Among the 13 knowledge questions examined, 10 exhibited significant differences between modes. In general, Web surveys elicit more accurate answers than face-to-face and CASI. Easy access to information through the Internet may contribute to the higher level of political knowledge among Web respondents. This finding suggests that political knowledge as measured in Web surveys is not equivalent to that from face-to-face surveys. The result suggests that previously established relationships between political knowledge and political involvement in face-to-face surveys may be different in Web surveys.

Introduction

Political knowledge is one of the most frequently measured variables in surveys and political polls. This variable has also been widely studied in political science research. For example, researchers have demonstrated the association between political knowledge and voting behavior, although the relationship is sometimes moderated by other factors, such as media consumption (Lanoue, 1992; Moore, 1987; Prior, 2005; Richey, 2008). Also, comparing and contrasting political knowledge across sub-groups, such as male and female, is another popular line of research (Dolan, 2011; Dow, 2009; Ondercin & Jones-White, 2011; Wolak & McDevitt, 2011).

Given the popularity and importance of political knowledge, it is surprising to see that limited attention has been devoted to the measurement of political knowledge questions, or knowledge questions in general. As pointed out by Nadeau and Niemi (1995), the cognitive process for answering knowledge questions is not as simple as retrieving relevant information or answering “don’t know” if such information is not readily available. In stead, people make educated guesses based on contextual cues and subjective attitudes. When it comes to measurement of political knowledge in particular, most of the studies by far have focused on the impact of the “Don’t know” option. For example, Mondak, (2001) found
that “Don’t know” and “Incorrect answer” imply two different concepts, although these two are often combined in analysis (see also Mondak, 1999). An explicit “Don’t know” option suppresses one’s willingness of expression and biases downward the actual knowledge level (Mondak & Davis, 2001). Also, the “Don’t know” option is a cause of gender inequality of political knowledge since men are more likely than women to guess when they are not certain (Mondak & Anderson, 2004). Studies by Miller and Orr (2008) and Robison (2014) confirmed Mondak and his colleagues’ findings by showing that respondents express more knowledge when the “Don’t know” selection is discouraged. However, Sturgis, Allum and Smith (2008) reported a counter-finding. In their study, respondents who initially provided a “Don’t know” answer were further probed to give a best guess. The best guess proved to be no better than chance, which indicates that the increased estimated knowledge level does not reflect their real knowledge level. More recently, Luskin and Bullock (2011) demonstrated that removing “Don’t know” only encourages random guess in close-ended questions but not in open-ended questions.

In this study, we expand the line of research on the measurement of political knowledge by examining the data collection mode effect on knowledge questions. Specifically, this study reports findings of the mode effect on political knowledge questions between face-to-face and Web surveys, and between computer-assisted self-interview (CASI) and Web surveys using national representative survey data. Although the mode effect among these three modes have been examined before, most of these investigations focus on response rate, data quality as measured by satisficing behaviors, and substantive responses to sensitive attitudinal and behavioral questions. Factual questions like political knowledge, which have true values, have received little attention from researchers.

Face-to-face versus web surveys

Three major findings can be drawn from mode studies on face-to-face versus Web surveys. First, face-to-face surveys lead to higher response rates than Web surveys do, possibly due to the higher level of interviewer contact in the former mode (Christensen, Ekholm, Glümer, & Juel, 2014; Heerwegh & Loosveldt, 2008). Second, Web surveys elicit more socially undesirable responses than face-to-face surveys do (Heerwegh, 2009). This is because of the lower level of interviewer involvement in the Web mode and the consequently higher level of anonymity. Third, the quality of data elucidated between these two modes is not conclusive. Although face-to-face surveys have lower item nonresponse rates and non-differentiation of rating scales, they suffer more from extreme response style bias than Web surveys do (Beukenhorst et al., 2014; Goldenbeld & de Craen, 2013; Heerwegh, 2009; Heerwegh & Loosveldt, 2008).

CASI versus Web survey

We are not aware of any study comparing these two modes directly. However, quite a few studies have compared one of these two self-administered modes with other self-administered modes, such as mail surveys (Brown, Vanable, & Eriksen, 2008; Halfors, Khatapoush, Kadushin, Watson, & Saxe, 2000; McCabe, Boyd, Couper, Crawford, & d’ Arcy, 2002). Previous studies have primarily focused on mode effects on sensitive questions and reached mixed findings. This is true for mode studies that involve CASI (Beebe, Harrison, McRae Jr, Anderson, & Fulkerson, 1998; Johnson et al., 2001; Webb, Zimet, Fortenberry, & Blythe, 1999; Wright, Aquilino, & Supple, 1998) or Web (Bälter, Bälter, Fondell, & Lagerros, 2005; Bates, Dahlhamer, & Singer, 2008; Denscombe, 2006; Knapp & Kirk, 2003; Link & Mokdad, 2005; McCabe et al.,
Mode effect on political knowledge

As far as we know, the study by Ansolabehere and Schaffner (2014) is the only one to examine the mode effect with regard to political knowledge. Their study compared three modes: national probability telephone surveys, mail surveys, and opt-in Web panels. The three knowledge questions concerned the unemployment rate, the party in control of the House of Representatives, and the party affiliation of the respondents’ states’ governors. The results revealed that Web respondents gave significantly more accurate answers to the unemployment rate question and less accurate answer to the house party question than mail respondents did. The party affiliation of governor question did not show a significant mode effect. Interestingly, after controlling for home Internet access, the mode effect on the three knowledge questions disappeared.

In the current study, we expand the line of study of the mode effect on political knowledge in two directions. First, we examine a broader range of political knowledge questions. Second, unlike the Ansolabehere and Schaffner (2014) study, the Web and face-to-face (including CASI module) surveys examined in this study are drawn from two independent national probability samples. Although coverage differences and self-selection bias may confound the result from the opt-in Web panel, this is not likely in our data.

Why do we expect there to be a mode effect for political knowledge questions? As suggested by Ansolabehere and Schaffner (2014), Web respondents can easily look up the answers on the Internet since they maintain the locus of control of the survey (Couper, 2011). Respondents in Web surveys can determine their flow and speed of response. Thus, pausing the survey and looking up the answers to the knowledge questions becomes a natural behavior. Respondents in face-to-face surveys, including the CASI module, are under a higher time constraint than Web survey participants are and it is less opportune for them to break the flow of the interview and look up the answers should they do not know them. Consequently, we expect more accurate answers in Web surveys than in face-to-face surveys and CASI.

Data and measures

The data we analyzed in this study comes from the 2012 American National Election Studies (ANES). The 2012 ANES conducted the survey in two modes: face-to-face and Web surveys, using two independent nationally representative samples. The face-to-face survey drew an address-based, stratified, multi-stage cluster sample. It contained a nationally representative main sample and oversamples for blacks and Hispanics. Intra-household selection was conducted and one eligible person was randomly chosen from each household. The face-to-face survey contained a CASI module for a subset of questions, including several political knowledge questions. The Web survey sample came from GfK KnowledgePanel, a nationally representative probability Web panel. The panel members were initially recruited through random-digit dialing or address-based sampling. Intra-household enumeration was performed during the recruitment stage and one person per household was selected to become the panel member. Selected households without Internet access were provided with free Internet service and hardware in order to
remove the coverage bias of the panel. As such, the panel was designed to be a representative of the
general U.S. population. (For more information about the GfK KnowledgePanel sample design, see
http://www.knowledgenetworks.com/ganp/docs/KnowledgePanel(R)-Design-Summary.pdf.) In the current
study, since the two probability samples both targeted the same population—U.S. citizens who were 18
years of age or older by the 2012 Election Day—they should produce comparable coverage.

Respondents in ANES completed two interviews: one pre-election and one post-election. The complete
pre-election survey contained 2054 face-to-face interviews and 3860 Web surveys. The response rates
(AAPOR RR1) were 38% and 2% for face-to-face and Web, respectively. The re-interview rates were 94%
and 93%, respectively. In the analysis throughout this study, we used the weight variable provided by the
survey organization, which accounts for the probability of household selection, the probability of
respondent selection within the household, nonresponse, and random sampling error. The weights were
post-stratified to produce estimates that match known population proportions for demographic
characteristics that are commonly used in post-survey adjustment. The post-stratification variables
include sex, age, race and ethnicity, educational attainment, metropolitan status, household Internet
access, income, marital status, home ownership, census region, and nation of birth. The post-
stratifications were conducted separately for face-to-face and Web surveys so that to adjustment both
samples to reflect the target population. In theory, the weighted analyses for both samples should both
produce the unbiased estimates for the same population. Therefore, the results were not likely to be the
consequence of different sample strategies or the differential nonresponse bias between the two modes.

In this study, we analyzed 13 knowledge questions from both the pre-election and post-election surveys.
The first two questions examined the religion of the Democratic (Obama) and Republican (Romney)
presidential candidates. The next five questions were asked in the CASI module of the face-to-face
survey. They ask about the number of times that one can be elected as the U.S. president, the present
federal budget deficit compared to the 1990s, the number of years which constitute one full term of office
for a U.S. Senator, the contents of Medicare and the area in which the U.S. federal government currently
spends the least. The above seven questions were in the pre-election survey. In the post-election survey,
there were four office recognition questions, which asked for the jobs or political offices that four people
hold in open-ended format: Joe Biden (Vice-President), David Cameron (Prime Minister of UK), John
Boehner (Speaker of the House), and John Roberts (U.S. Supreme Court Chief Justice). The last two
questions ask about the party that had the most members in the House of Representatives and the U.S.
Senate. In the analysis, we compared the correct answers between face-to-face versus Web surveys and
CASI versus Web surveys. For the four office recognition questions, the answers were coded as being
either correct or incorrect by the survey organization and we used this in the analysis. We computed the
mean for the corrected correct answers for each mode. Then, we compared the means between modes
using an independent t-test. All analyses were weighted so that their results reflected the survey
population.

Results

In this section, the differences in demographic compositions between the face-to-face and Web surveys
are reported. As can be seen from Table 1, demographic variables, including gender, age, race and
ethnicity, education level, marital status and household income are not significantly different between the
two modes. In both modes, the respondents tended to be non-Hispanic whites, with high school or some
college education, married, and making $49,999 or less household-wide. Gender and age were more
evenly distributed. This result indicates that there is not likely to be a differential nonresponse bias between the modes. Thus, we are confident that differences between modes, if any, are not attributable to sample composition differences.

Table 1. Means and Standard Errors (S.E.) for Demographic Variables by Mode of Data Collection, 2012 American National Election Studies.

<table>
<thead>
<tr>
<th></th>
<th>Face-to-face</th>
<th>Web</th>
<th>Chi-square</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.E.</td>
<td>Mean</td>
<td>S.E.</td>
</tr>
<tr>
<td>Female</td>
<td>0.52</td>
<td>0.02</td>
<td>0.52</td>
<td>0.01</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>0.21</td>
<td>0.01</td>
<td>0.21</td>
<td>0.01</td>
</tr>
<tr>
<td>30~39</td>
<td>0.16</td>
<td>0.01</td>
<td>0.15</td>
<td>0.01</td>
</tr>
<tr>
<td>40~49</td>
<td>0.17</td>
<td>0.01</td>
<td>0.17</td>
<td>0.01</td>
</tr>
<tr>
<td>50~59</td>
<td>0.19</td>
<td>0.01</td>
<td>0.20</td>
<td>0.01</td>
</tr>
<tr>
<td>60~69</td>
<td>0.14</td>
<td>0.01</td>
<td>0.16</td>
<td>0.01</td>
</tr>
<tr>
<td>70+</td>
<td>0.13</td>
<td>0.01</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>Race and ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White (non-Hispanic)</td>
<td>0.71</td>
<td>0.01</td>
<td>0.71</td>
<td>0.01</td>
</tr>
<tr>
<td>Black (non-Hispanic)</td>
<td>0.12</td>
<td>0.01</td>
<td>0.12</td>
<td>0.01</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.11</td>
<td>0.01</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>Other (non-Hispanic)</td>
<td>0.06</td>
<td>0.01</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school credential</td>
<td>0.10</td>
<td>0.01</td>
<td>0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>High school credential</td>
<td>0.30</td>
<td>0.01</td>
<td>0.30</td>
<td>0.01</td>
</tr>
<tr>
<td>Some post-high-school</td>
<td>0.30</td>
<td>0.01</td>
<td>0.30</td>
<td>0.01</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>0.19</td>
<td>0.01</td>
<td>0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>Graduate degree</td>
<td>0.10</td>
<td>0.01</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.53</td>
<td>0.02</td>
<td>0.53</td>
<td>0.01</td>
</tr>
<tr>
<td>Widowed</td>
<td>0.06</td>
<td>0.01</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Divorced</td>
<td>0.13</td>
<td>0.01</td>
<td>0.13</td>
<td>0.01</td>
</tr>
<tr>
<td>Separated</td>
<td>0.03</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table 2 reports the proportions of correct answers to the knowledge questions in different modes. Column 1 contains the proportions in Web surveys, and columns 2 and 3 contain the proportion in the face-to-face survey and CASI module from all respondents and respondents with Internet access, respectively. The first two questions about the religion of the Democratic and Republican Presidential candidates do not show significant differences between modes. The four questions on office recognition, including questions about Joe Biden, David Cameron, John Boehner, and John Roberts, and the two questions on the party with the most members in the House and Senate, all displayed a significant mode effect but the directions are different. For the office recognition questions, Web survey respondents consistently provided more correct answers than face-to-face survey respondents did, regardless of Internet accessibility. The only exception to this is the Joe Biden office recognition question, in which no significant difference was detected between Web respondents and face-to-face respondents who had Internet access. The trend reversed for the two questions about which party won the most seats in the House and Senate. For these two questions, face-to-face respondents provided more correct answers than Web respondents did. The result remains almost identical after restricting the comparison to face-to-face respondents with Internet access.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(1)-(2)</th>
<th>(1)-(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Web</td>
<td>Face-to-face All</td>
<td>Face-to-face with Internet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democratic Presidential candidate religion</td>
<td>33</td>
<td>37</td>
<td>37</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>Republican Presidential candidate religion</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>Joe Biden</td>
<td>88</td>
<td>84</td>
<td>86</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>David Cameron</td>
<td>24</td>
<td>10</td>
<td>11</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>John Boehner</td>
<td>45</td>
<td>31</td>
<td>32</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>John Roberts</td>
<td>39</td>
<td>21</td>
<td>22</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>House most party before election</td>
<td>71</td>
<td>76</td>
<td>76</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>Senate most party before election</td>
<td>65</td>
<td>71</td>
<td>70</td>
<td>-6</td>
<td>-5</td>
</tr>
<tr>
<td></td>
<td>Web</td>
<td>CASI All</td>
<td>CASI with Internet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Times President can be elected</td>
<td>93</td>
<td>87</td>
<td>89</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Size of federal deficit: Bigger</td>
<td>87</td>
<td>82</td>
<td>83</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
When comparing the five questions in the CASI module of the face-to-face survey with the Web survey, all questions except for the one asking which area the federal government spent the least on exhibit significant mode effect. The other four questions, including the number of times that one can be elected as the President, size of the federal budget deficit compared to the 1990s, the length of one term as a Senator, and the contents of Medicare, all showed a significant mode effect. In particular, Web survey respondents consistently provided more accurate answers than respondents to the CASI module in the face-to-face survey did. This is true regardless of the Internet availability of the face-to-face sample.

In addition to the above analysis, we also examined the mode effect by two subgroups—respondents’ education attainment and country of birth. Specifically, we compared the mode effects for people with education level of high school or less versus more than high school, and for people who were born in the U.S. versus those who were born in other countries. The results (not shown, available from the author upon request) for each of the subgroup showed the same pattern as the result from the whole sample reported above and the general conclusion remained unchanged.

**Discussion**

Out of 13 political knowledge questions in the 2012 American National Election Studies, 10 questions manifested a significant mode effect. Respondents in the Web survey gave more accurate answers to eight political knowledge questions than the face-to-face survey or the CASI module yielded. When comparing face-to-face versus Web surveys, only the two questions about the religions of presidential candidates were not susceptible to the mode effect. For the CASI versus Web survey comparison, answers to four out of the five questions were significantly different across modes. The only exception was the question on Federal government spending, although the accuracy rates are low in both modes. After restricting the comparison to the Web versus face-to-face respondents with Internet access at home, the results remained almost the same. The two questions asking about the dominant party in the House and Senate received more accurate face-to-face responses than the Web survey elicited. One possible explanation for this is that compared to the other knowledge questions, the answers to these two questions were more difficult to find on the Internet since the key words are not immediately clear. The relatively more complex searching task may have contributed to the lower level of accuracy in the Web survey.

The results from this study have important implications for measuring the general population’s political knowledge. Traditionally, knowledge questions have been designed to measure how much the public knows about domestic and international political affairs and were conducted through face-to-face and telephone surveys. In Web surveys, knowledge questions do not measure political knowledge in the traditional sense; that is, the amount of knowledge that people carry around. Rather, respondents can access information immediately if they do not already have it. Therefore, researchers need to re-think the
meaning of political knowledge in Web surveys. In the era of the Web survey, the concept of political knowledge may need to be redefined. In that case, some of the established relationships between political knowledge and behaviors will become questionable. For example, if Web respondents provide accurate answers to knowledge questions only through searching the answer on the spot during the survey, then this accuracy may not indicate a higher level of political interest and involvement. Therefore, the predictability of voter turnout from political knowledge may not be as reliable in Web surveys as it is in face-to-face surveys.

Like any other research, this study also has its own limitations that need to be addressed in future research. First, although we suspect that the higher level of political knowledge is the result of easy accessibility to information of the Web respondents, we do not have data to support such argument. Future study should collect the relevant paradata, such as response latency and the numbers of intermediate clicks to the political knowledge questions. A longer response time, especially after change of answer from incorrect to correct one, would be a strong indication of looking up the answers online. Second, future study should also use experimental design to replicate this study. Although we think the coverage is comparable between the two samples in this study and the weighted analysis should remove any potential nonresponse bias, it will still be a useful effort to randomly assign respondents from one sample source to one of the two modes.

Our research only paints an initial picture of the mode effect on the political knowledge. Given the wide usage and importance of this measure, it is clear that more work needs to be done to further explore the survey methodological aspect of political knowledge questions.

Appendix. Question wordings and response options. [Notes in brackets are added by the authors.]

Now we would like to ask you some questions about the religion of the presidential candidates. Would you say that [Obama] is Protestant, Catholic, Jewish, Muslim, Mormon, some other religion, or is he not religious? [Pre-election survey, face-to-face versus Web. Correct answer: Protestant]

Would you say that [Romney] is Protestant, Catholic, Jewish, Muslim, Mormon, some other religion, or is he not religious? [Pre-election survey, face-to-face versus Web. Correct answer: Mormon]

Do you happen to know how many times an individual can be elected President of the United States under current laws? [Pre-election survey, CASI versus Web. Correct answer: twice]

Is the U.S. federal budget deficit – the amount by which the government’s spending exceeds the amount
of money it collects – now bigger, about the same, or smaller than it was during most of the 1990s? [Pre-election survey, CASI versus Web, open-ended question. Correct answer: bigger]

For how many years is a United States Senator elected – that is, how many years are there in one full term of office for a U.S. Senator? [Pre-election survey, CASI versus Web, open-ended question. Correct answer: 6 years]

What is Medicare? [Pre-election survey, CASI versus Web]

1. A program run by the U.S. federal government to pay for old people’s health care [Correct answer]
2. A program run by state governments to provide health care to poor people
3. A private health insurance plan sold to individuals in all 50 states
4. A private, non-profit organization that runs free health clinics

On which of the following does the U.S. federal government currently spend the least? [Pre-election survey, CASI versus Web]

1. Foreign aid [Correct answer]
2. Medicare
3. National defense
4. Social Security

Now we have a set of questions concerning various public figures. We want to see how much information about them gets out to the public from television, newspapers and the like. [Post-election survey, face-to-face versus Web]

The first name is: John Boehner. What job or political office does he now hold? [Correct answer: Speaker of the House]

Joe Biden. What job or political office does he now hold? [Correct answer: Vice-President]
David Cameron. What job or political office does he now hold? [Correct answer: Prime Minister of UK]

John Roberts. What job or political office does he now hold? [Correct answer: U.S. Supreme Court Chief Justice]

Do you happen to know which party had the most members in the House of Representatives in Washington before the election [this/last] month? [Post-election survey, face-to-face versus Web]

1. Democrats

2. Republicans [Correct answer]

Do you happen to know which party had the most members in the U.S. Senate before the election [this/last] month?

1. Democrats [Correct answer]

2. Republicans

References


