

**The Internet as a Public Opinion Laboratory:
Experiments with Survey Questions***

George Bishop and B.J. Jabbari

Internet Public Opinion Laboratory

**Department of Political Science
University of Cincinnati**

***Paper presented at the annual conference of the American Association for
Public Opinion Research, Montreal Canada, May 17-20, 2001.**

The Internet as a Public Opinion Laboratory

Surveys on the World Wide Web have become ubiquitous and will continue to proliferate, whether we like it or not, now that the technological genie is out of the bottle. Nearly everyone with access to the Internet, it seems, can find a way to mount a poll. Just recently, Mick Couper (2000) has critically reviewed the various types of Web Surveys that have come into being and discussed the issues of survey quality—coverage and sampling error, nonresponse error, and measurement error—that need to be addressed relative to the low cost and speed with which Internet surveys can be executed (see also Black, 1998; Mitofsky, 1999). The challenge, as he sees it, is how to design Web surveys that maximize the potential of this new data collection medium while minimizing error. We will not address these issues here, but focus instead on the potential of using the Internet as a public opinion laboratory for conducting experimental research and hypothesis testing with large and sufficiently diverse samples of Internet users. We view the Internet not only as an inexpensive, high-speed alternative to using conventional campus laboratory experiments with college sophomores, but also as a more effective vehicle for experimental testing than large-scale RDD surveys. Taking a line from Berkowitz and Donnerstein (1982), we also believe that external validity is “more than skin deep”: though representative sample designs are certainly essential for estimating population parameters, they are not necessarily required for making causal inferences.

Making causal inferences about the effects of survey questions is a conspicuous example. Ever since Gallup and his peers began polling in the mid-1930s, public opinion researchers have conducted “split-ballot” experiment (as many of us still call them) to determine the effects of variations in the way questions are worded, the format in which

they are presented, and the order and context in which they are asked (see Bishop and Smith, 1991, Rugg and Cantril, 1944). In what is now considered as a classic set of such experiments, Schuman and Presser (1981) have probed the complex effects of question order and response order; of open vs. closed questions; of offering or not offering respondents an explicit opportunity to say “no opinion”; of including or omitting a middle response alternative; of asking balanced or imbalanced questions—in an agree-disagree or forced-choice format; and of wording questions with a biased or leading tone. All of these experiments were carried out in large-scale, national probability surveys with therefore a high degree of generalizability. So too have the vast majority of other split-ballots in the literature been carried out with large national or regional probability samples of the adult population (see Moore and Newport, 1996, for some recent examples). All of this is well and good, but we think much the same can now be done more efficiently and effectively using the Internet as a public opinion laboratory. We carried out just such an Internet experiment this past fall.

Cincinnati Web Experiment

The 2000 presidential election campaign provided the context and opportunity to conduct our first Internet experiment with survey questions. During the final month of the campaign (October 8-November 7) we created a browser-based Web survey of the candidates and the issues. The respondents to the survey constituted an entirely self-selected, nonprobability sample of various and sundry populations: students at the University of Cincinnati who were recruited through notices to faculty teaching classes in computer science, political science, and other Arts & Sciences departments; visitors to the University's home page where we had a banner ad and link; visitors to a local

television station's web site (Channel 9) where we also had a link; visitors to the websites of two local weekly newspapers (City Beat, Citizen); local and regional websites of various political parties; and volunteers from who knows where else, including AAPORITES who became aware of the survey through an announcement on AAPORNET. Our aim was to generate a sufficiently large and demographically diverse set of respondents solely for purposes of experimentation, replication, and causal inference. We had no illusions about weighting the data and no weighting was done.

Table 1 shows a demographic profile of this self-selected sample of volunteers (N=1356). As compared to the general population, and perhaps even to the user population on the Internet, this group of respondents was much younger; better educated; more likely to be male, white, nonreligious, and registered to vote; and identified with other political parties—reflecting, in significant part, the strong participation by students from the University of Cincinnati and the surrounding community. So it is a highly self-selected set of respondents and yet *sufficiently varied* for purposes of hypothesis testing and replication. And, by definition, the response rate for any such self-selected sample is unknown.

Research Design

The importance, first and foremost, of being able to replicate well-known types of question effects in the literature drove our design (see, e.g., Schuman and Presser, 1981). For this initial outing, we chose to design a set of split-sample experiments on the effects of: offering vs. not offering respondents an explicit *don't know*, *not sure*, or *no opinion* response alternative; including or omitting a middle response alternative; response order; and tone of wording (forbidding vs. allowing prayer and the Ten

Commandments in public schools). To randomly assign respondents to alternate versions of the questionnaire we used an **odd-or-even second method**: visitors who landed on our website on an odd-numbered second received Form A of the questionnaire; those who landed on an even-numbered second got Form B. Most of the questions we asked focused on issues in the 2000 election campaign that we adapted for self-administration on the Web from Gallup and other polling organizations and thus represent conceptual rather than exact replications of question effects in the literature. The first phase of the experiment ran from October 8-24; the second phase, in which we modified some of the experimental conditions, ran from October 24-November 7 (election day).

The DK Experiments. In this presentation we will have time to look at just one subset of these experiments: the effects of offering or not offering respondents an explicit opportunity to choose *don't know*, *not sure*, or *no opinion* as an alternative. The standard practice of most polling organizations is to make such responses strictly voluntary and thus minimize missing data. Some organizations prefer to use "not sure"; others favor DK or "no opinion." In the first phase of our initial experiment respondents received either a completely *forced-choice form* in which no DK option at all was available or a *filtered form* in which three such options were available: Not Sure, Don't Know, or No Opinion. In the second phase respondents were given either the same filtered form or a modified forced-choice form, which included an "Other" category in an attempt to simulate a volunteered response. These experimental variations were used for six social and economic issues: social security investment in private stocks and bonds, the long-term effects of such investment, whether to allow local governments to sue gun manufacturers, managed health care, the death penalty, and abortion. In addition, we embedded a

question about a fictitious *1995 Social Security Reform Act* immediately after the questions on social security investment in the market. Here we will examine the data for only the questions on social security, though the pattern of results appears to hold for all the other issues we included in the DK experiments. The social security questions, which were identical to those asked by the Gallup Organization in its national polls during campaign 2000, read as follows:

1. "Do you favor or oppose a proposal that would allow people to put a portion of their Social Security payroll taxes into personal retirement accounts that would be invested in private stocks and bonds?"

2. "Thinking about the long-term condition of the Social Security system, if people were allowed to put a portion of their Social Security payroll taxes into accounts that would be invested in the stock market, do you think this would strengthen or weaken the ability to pay benefits to retirees 25 years from now?"

Findings

The figures for Experiment I in table 2 tell a familiar tale. Given an explicit opportunity, many respondents will avoid giving an opinion by indicating they are *not sure*, *don't know*, or have *no opinion*. In this case nearly one out of five (fairly well-educated) respondents (18.2%) indicated they did not have an opinion on the issue of investing social security payroll taxes in the market. And more than one out of four (27.6%) lacked an opinion on the related question of whether investing such taxes in the market would strengthen or weaken the ability to pay benefits to retirees 25 years from

now. Notice too that, by a margin of 2 or 3-1, respondents generally preferred *not sure* as an option over either *don't know* or *no opinion*. This pattern suggests that on this and other issues (not shown here) survey respondents may be more uncertain about which way they lean on the issue—a middle type of response—than lacking in the knowledge to express an opinion, a nuance that allows us to separate distinct components of the typically undifferentiated DK response in public opinion surveys.

The results for the forced-choice form in Experiment I tell us nothing directly about the DK response since respondents were, by definition, unable to volunteer a “don't know,” as would normally be the case in a phone interview or by writing it in on a self-administered mail questionnaire, for example.

To simulate an ambiguous, but explicit response option, we modified the forced-choice form in Experiment II to include an “Other” category. The results in table 2 show that, while this option obviously relieves the pressure to give a response—versus letting respondents skip the question and thereby generate unwanted missing data—it adds no useful information as compared to giving them an explicit chance to indicate *Not Sure*, *Don't Know*, or *No Opinion*. As in Experiment I the percentage of respondents selecting one of these three options was fairly similar in magnitude and in roughly the same ratio, with *not sure* responses chosen noticeably more often than either *don't know* or *no opinion*.

Effects on the Distribution of Opinion

But does any of this make a difference in the conclusions a researcher would draw about the distribution of public opinion on a given issue? Previous experiments in RDD surveys with national and regional samples have shown that, while offering

respondents an explicit DK response alternative will typically increase the magnitude of such responses by 20-25%, it generally makes little or no difference in one's conclusions once the DK responses are excluded from the analysis (Schuman and Presser, 1981; see also Bishop et al., 1983 for replications as well as a notable exception). The results in table 3 (and those for other issues not shown here) replicate these findings exactly, and at much less expense, with a highly self-selected, nonprobability sample of respondents to our Web survey. With DK responses excluded from the analysis we find highly similar percentages of respondents on both the forced-choice and filtered forms of the questions in favor of allowing people to invest their social security taxes in the market (76.2% vs. 77.7%, respectively), and nearly identical percentages of respondents indicating that they believe putting social security taxes in the market would strengthen, rather than weaken, the ability to pay benefits to retirees 25 years from now (72.9% vs. 72.3%, respectively). Conclusions about the distribution of opinion would thus be the same regardless of which form was used.

Effects on Relationships between Opinions

The exact same pattern of replication occurs when we examine the association between responses to the two questions on social security. The vast majority of those respondents who favored investing social security taxes in the market also believed it would strengthen, rather than weaken, the ability to pay retirees 25 years from now, and vice versa for those who opposed putting such funds in the market. The magnitude of association was extremely high and nearly identical with both the forced-choice and filtered forms of the questions, with no evidence whatsoever of any interaction effect (table 4). One would again draw the same conclusion—in this case about the relationship

between these opinions—irrespective of which question form was used. Much the same holds for other issues we examined in these experiments—the relation between opinions on gun control and the death penalty, opinions on abortion and political ideology, and opinions on managed health care and ideology (data not shown here): the form of the question makes little or no difference in the size of the associations; they appear to be “form-resistant” as it is sometimes described (cf. Schuman and Presser, 1981, Ch. 4). Once more, we have a conceptual replication of experiments from the response effects literature at a fraction of the cost of a large-scale, national or regional sample survey.

Opinions on a Fictitious Issue

The questions on social security included one other conceptual replication (Bishop et al., 1980): a question about a nonexistent issue. Following the two questions about investing social security taxes in the market, respondents were asked: “Do you favor or oppose a proposal to repeal the 1995 Social Security Reform Act?” On Form A, to which half the respondents were randomly assigned, the response categories were: Favor, Oppose, or Haven’t Heard of it. On Form B, to which the other half was assigned, the options were: Favor, Oppose, Not Sure, Don’t Know, and No Opinion. So on both versions of the question respondents were given an explicit opportunity to say they had not heard of it or did not have an opinion on it in one way or another.

But despite this encouragement to choose some form of DK response, a remarkably high percentage of respondents offered an opinion on repealing the Social Security Reform Act: about 1 out of 4 respondents (25%) on Form A and a little over 1 out of 5 (22%) on Form B (table 5). These figures look all the more remarkable because the respondents to this Web survey were much better educated than the average adult, a

factor which should have reduced substantially the percentage of respondents volunteering an opinion given what we knew from previous surveys on fictitious or obscure issues, where the range has typically been about 25-30% (see e.g., Schuman and Presser, 1981, Ch. 5).

The relationship of education to offering an opinion on the Social Security Reform Act also varied in an unusual way with the question form (table 6). On Form A the percentage giving an opinion did not vary significantly by education: about three-fourths of the respondents indicated they had not heard of it, regardless of education. But on Form B, and contrary to previous research findings (Bishop et al., 1980; Schuman and Presser, 1981), Highly educated respondents were somewhat more likely than the less well educated to volunteer an opinion on repealing the Social Security Reform Act (and more opposed to repealing it for reasons that escape us). The highly educated were perhaps somewhat more likely to think they knew something about the topic from having answered the previous questions about investing social security funds in the market; they were, in fact, more likely than the less well educated to have expressed an opinion on the immediately preceding question about whether investing such funds in the market would strengthen or weaken the ability to pay benefits to retirees 25 years from now (data not shown here). So while these findings fail to replicate the usual pattern in previous experiments, they suggest an alternative explanation about how respondents, regardless of education, impute meaning to ambiguous survey questions from the immediately preceding context—a psychological situation not unlike the one we find in current national polls on President Bush's proposal to reform the social security system, and an hypothesis readily pursued at much less expense through the laboratory of the Internet.

Conclusion

You don't need a national probability sample to test a hypothesis. Whether it's just about testing the effects of different question forms or contexts or a causal model of public opinion and political behavior, a nonprobability sample of Web users will probably do. It is likely to be much larger and more diverse than the typical college sophomore sample of social-psychological research and can be obtained and replicated far more rapidly. The WEB questionnaires themselves can readily incorporate elaborate randomized designs to test complex, cognitive models of survey question effects, and in a variety of multi-media. Here we have started with a simple replication of a few classic experiments from the literature to illustrate the enormous potential of the Internet as a public opinion laboratory. We've only just begun.

References

- Berkowitz, Leonard, and Edward Donnerstein. 1982. "External Validity is More Than Skin Deep." American Psychologist 37: 245-257.
- Bishop, George, Robert W. Oldendick, and Alfred J. Tuchfarber. 1983. "Effects of Filter Questions in Public Opinion Surveys." Public Opinion Quarterly 47: 528-546.
- Bishop, George F., Robert W. Oldendick, Alfred J. Tuchfarber, and Stephen E. Bennett 1980. Pseudo-Opinions on Public Affairs." Public Opinion Quarterly 44: 198-209.
- Bishop, George, and Andrew Smith. 1991. "The Gallup Split-Ballot Experiments." Public Perspective 2: 25-27.
- Black, Gordon S. 1998. "The Advent of Internet Research." Paper presented at the annual conference of the American Association for Public Opinion Research. St. Louis, May.
- Couper, Mick P. 2000. "Web Surveys: A Review of Issues and Applications." Public Opinion Quarterly 64: 464-494.

Mitofsky, Warren. 1999. "Pollsters.Com." Public Perspective June/July: 24-26.

Moore, David, and Frank Newport. 1996. "Public Policy Questions and Response Order: Prevalence of the Recency Effect." Paper presented at the annual meeting of the American Association for Public Opinion Research, Salt Lake City, Utah, May.

Rugg, D., and Hadley Cantril 1944. "The Wording of Questions." In H. Cantril (ed.), Gauging Public Opinion. Princeton University Press.

Schuman, Howard, and Stanley Presser. 1981. Questions and Answers in Attitude Surveys. Academic Press.

Table 1. Demographic and Political Profile of Respondents in the Cincinnati Web Experiment

Age Group

Under 25	38.4%
25-29	13.3
30-39	15.9
40-49	13.9
50-59	12.1
60 or older	6.4

Race

White	88.1%
Black	4.2
Hispanic	0.6
Asian	3.2
Multi-Racial	0.8
Other	3.1

Sex

Male	62.8%
Female	37.2

Education

High School or less	13.2%
Some College	34.1
College Graduate	22.1
Postgraduate Study	30.6

Religion

Catholic	29.0%
Protestant	33.0
Jewish	4.4
Other	13.0
None	20.6

Registered to Vote

Yes	87.0%
No	13.0

Table 1(continued)

Party Identification

Democrat	29.3%
Republican	40.9
Green	2.1
Libertarian	2.9
Reform	0.3
Natural Law	0.2
Other Party	1.3
Independent	23.1

Source: *Presidential Election Poll: Campaign 2000*. Internet Public Opinion Laboratory, Department of Political Science, University of Cincinnati.

Table 2. Experiments with Social Security Issues

Experiment I

Forced-Choice Form

Filtered Form

Investment of Social Security Taxes

Do you favor or oppose a proposal that would allow people to put a portion of their social security payroll taxes into personal retirement accounts that would be invested in private stocks and bonds?

Favor	76.2%
Oppose	23.8

	100.0
	(462)

Do you favor or oppose a proposal that would allow people to put a portion of their social security payroll taxes into personal retirement accounts that would be invested in private stocks and bonds?

Favor	63.6%
Oppose	18.2
Not Sure	11.8
Don't Know	3.3
No Opinion	3.1

	100.0
	(483)

Strengthen or Weaken Social Security

Thinking about the long-term condition of the Social Security system, if people were allowed to put a portion of their Social Security payroll taxes into accounts that would be invested in the stock market, do you think this would strengthen or weaken the ability to pay benefits to retirees 25 years from now?

Strengthen	72.9%
Weaken	27.1

	100.0
	(462)

Thinking about the long-term condition of the Social Security system, if people were allowed to put a portion of their Social Security payroll taxes into accounts that would be invested in the stock market, do you think this would strengthen or weaken the ability to pay benefits to retirees 25 years from now?

Strengthen	52.4%
Weaken	20.1
Not Sure	20.1
Don't Know	4.8
No Opinion	2.7

	100.0
	(483)

Table 3. Opinion on Social Security Issues by Question Form (DKs Excluded)

Experiment I

Forced-Choice Form

Filtered Form

Investment of Social Security Taxes

Favor	76.2%	77.7%
Oppose	23.8	22.3
	-----	-----
	100.0	100.0
	(462)	(395)

$$\chi^2 = .28, df=1, n.s.$$

* Omits 88 observations

Strengthen or Weaken Social Security?

Strengthen	72.9%	72.3%
Weaken	27.1	27.7
	-----	-----
	100.0	100.0
	(462)	(350)

$$\chi^2 = .04, df=1, n.s.$$

* Omits 133 observations

Table 4. Response to Strengthen or Weaken Social Security by Investment of Social Security Taxes, by Question Form

Experiment I

Forced-Choice Form

	<u>Investment of Social Security Taxes</u>	
	Favor	Oppose
Strengthen	91.5%	13.6%
Weaken	8.5	86.4
	-----	-----
	100.0	100.0
	(352)	(110)
$\chi^2 = 257.31, df=1, p<.001$ $\gamma = .97$		

Filtered Form

	<u>Investment of Social Security Taxes</u>	
	Favor	Oppose
Strengthen	94.4%	7.4%
Weaken	5.6	92.6
	-----	-----
	100.0	100.0
	(249)	(68)
$\chi^2 = 219.9, df=1, p<.001$ $\gamma = .99$		

Strengthen or Weaken Social Security x Investment of Social Security Taxes x Form:

$$\chi^2 = 3.43, df=1, n.s.$$

Table 5. Responses to Question about Repealing The 1995 Social Security Reform Act by Question Form

	Form A		Form B
Favor	13.1%	Favor	12.5%
Oppose	11.7	Oppose	9.4
Have Not Heard of It	<u>75.2</u>	Not Sure	25.7
	100%	<i>Don't Know</i>	38.9
	(N=674)	No Opinion	<u>13.6</u>
			100.1%
			(N=682)

Table 6. Response to 1995 Social Security Reform Act by Education, by Question Form

Form A

	High School or Less	Some College	College Grad	Postgraduate Study
Favor	14.8%	13.6%	17.9%	8.3%
Oppose	8.6	10.6	9.3	16.0
Have Not Heard	<u>76.5</u>	<u>75.8</u>	<u>72.8</u>	<u>75.7</u>
(N=)	99.9 (81)	100.0 (236)	100.0 (110)	100.0 (156)

Form B

	High School or Less	Some College	College Grad	Postgraduate Study
Favor	11.2%	11.5%	17.6%	10.5%
Oppose	8.2	6.2	4.7	16.7
Not Sure	29.6	31.7	19.6	21.5
<i>Don't Know</i>	35.7	36.6	40.5	41.6
No Opinion	<u>15.3</u>	<u>14.1</u>	<u>17.6</u>	<u>9.6</u>
	100.0 (98)	100.1 (227)	100.0 (148)	99.9 (209)