

Comparing the Knowledge Networks Web-Enabled Panel
and the In-Person 2002 General Social Survey:
Experiments with Mode, Format, and Question Wordings

Tom W. Smith

National Opinion Research Center/University of Chicago

J. Michael Dennis

Knowledge Networks

March, 2004

GSS Methodological Report No. 99

1.1

The authors wish to thank Richard Li of Knowledge Networks (KN) for his direction of the KN data collection and his assistance with analysis.

Introduction

Social scientists and others have become increasingly interested in using email and the Internet to conduct surveys. Major factors impede the general and widespread use of e-surveys including limited population coverage, differential non-coverage, the difficulty of conducting random or probability sampling, low response rates, etc. But other factors make e-surveys attractive, including low costs, the use of the self-administered mode, and the ability to incorporate visual images.¹ Moreover, it is unclear how comparable e-survey are to those utilizing more traditional modes (in-person, mail, and telephone) even when issues of coverage and probability sampling can be dealt with. Little empirical evidence has been collected on how the measurement features and error structures of e-surveys compare to those of other modes.

The Experiments

To examine the similarities and differences between web-based and in-person surveys, experiments were designed to compare the 2002 General Social Survey (GSS) and a 2002 study by Knowledge Networks (KN). These experiments partly replicate and extend experiments done on the 2000 GSS and a 2000 KN study (Smith, 2003).

¹On the different types of e-surveys that exist and the general features and performance of same see Bosnjak, Tuten, and Bandilla, 2001; Couper, 2001; Dillman, 2000; and Smith, 2003.

The 2002 GSS is an in-person, multi-stage, area probability sample of adults living in households in the United States conducted by the National Opinion Research Center (NORC), University of Chicago. The 2002 GSS had a sample size of 2765 and a response rate of 70.1% (Response rate # 3 of the American Association for Public Opinion Research (AAPOR)). The field dates were from early February to June. For more details see Davis, Smith, and Marsden, 2003. The Knowledge Networks survey (KNS) was conducted on their web-enabled panel. KN respondents consist of people contacted via RDD surveys (originally conducted by NORC and currently by KN itself) and recruited into a panel. Recruits are provided with WebTV and access to the Internet via WebTV or use their own personal computer to participate. In exchange for this and other incentives the panelists agree to periodically answer surveys sent to them by email. KNS had a total sample size of 1655 and a response rate of about 13%, taking into consideration each stage of response (AAPOR # 3). Data collection was April-June, 2002. General information on KN can be found at www.knowledgenetworks.com/ganp.

The questions compared in the experiments are known as the spending priority items. They were the first items on both the 2002 GSS and KNS (thereby controlling for context). The wordings are given in Table 1. The 2002 GSS, like the 2000 GSS, asked the 11 standard wordings and 11 variant wordings on random sub-samples. These were then each followed by the 6 spending items without variant wordings. This same between subjects design was used on the

2000 KNS, but in 2002 KNS used a within subjects design in which the 11 standard items were asked first followed by the other 6 standard items and then the 11 variant items.

The 2002 GSS and the KNS experiments allow the following comparisons: 1) in-person, GSS to KNS panel, 2) Don't Know (DK) formats within KNS and each to the GSS format, 3) question wording experiments within GSS and KNS and across modes, and 4) 2002 results to similar mode and wording experiments, but not DK format experiments, in 2000 between the GSS and KNS.

Don't Know Levels across Formats and Modes

Table 1 show the DK levels in the 2002 GSS and the four KNS treatments. On the GSS Don't Know is an unread response. A special key (F8) is used by the interviewer to capture such responses when respondents give such a reply. The KNS versions were administered to random sub-samples and consist of the following formats:

A = Instruction on first screen to skip questions to indicate "Don't Know/No Opinion". It read "In this survey we'd like your opinions about some important national issues. There is no right or wrong answer and please answer the questions based on your knowledge. If you do not know the answer to a question, you can simply skip the question to indicate 'Don't know' or 'No opinion'." Don't Know option is not presented on screen as a response option for individual items.

B = Don't Know presented as a response option for each individual item.

C = Don't Know option presented on screen as a response option for each individual item. Then respondents who initially selected Don't Know are prompted to select a response on a follow-up screen that asked, "If you had to choose, are we spending too much money, too little money, or about the right amount of money on..."

D = Instructions on all screens "If you do not know the answer to the question, you can simply skip the question to indicate 'Don't Know' or 'No Opinion'." Skipping is done by clicking on a "Next Question" button. Don't Know not listed among response options.

Table 2 averages and summarizes results from Table 1. It shows that the Don't Know levels average the lowest for the KNS-A treatment (with skip-out instructions on the initial screen only) (3.2% for the first 11 standard items), followed by the GSS (3.4%), KNS-C (with explicit Don't Know among the response options followed by a probe of all DK answers) (10.6%), KNS-D (skip-out instructions on each screen) (11.0%), and KNS-B (with explicit Don't Know among the response options on each screen and no follow-up probes (16.2%). Clearly Don't Knows are mentioned more frequently when they are made more explicit and acceptable. This is consistent

with the literature on format-variable DK levels (Schuman and Presser, 1981; Smith, 1982; Smith, 1985). This pattern also appears within the GSS between 2000 and 2002 when the GSS shifted from paper-and-paper interviewing (PAPI) to computer-assisted-personal interviewing (CAPI). The % mentioning DKs declined in 2002 presumably because CAPI uses more of a skip-out approach while PAPI has DKs explicit among the response options (Smith and Kim, 2003). Across 49 opinion items DKs averaged 4.4% on the 1998 and 2000 GSSs and 2.6% on the 2002 GSS. This indicates that interviewers as well as respondents are influenced by the format of DKs in questions.

Table 2 also shows a within-format, learning effect. On the KNS within-subjects design people were first asked the 11 standard questions with parallel variant items, next the 6 other standard items without variant versions, and then the 11 variant wordings to the first 11 items. On GSS and KNS-A, B, and D DKs are lowest for the first 11 items, a bit higher for the 6 other standard items, and intermediate for the 11 variant items. But on KNS-C with the follow-up prompt to those initially saying Don't Know, the DK level before the probe falls by over 50% from the initial 11 items to the 6 other standard items (from 10.6% to 5.0%) and is about the same for the 11 variant items (4.8%). This strongly suggests a learning effect in which people reduced mentions of DKs to avoid the follow-up prompts once they realized that their initial DK responses would be probed. This is further supported by looking at the DK levels on KNS-D and KNS-B. For the no-probe version (KNS-D) DKs for the first 6 standard items averaged 10.9% and for the last five items

averaged 11.1%, For KNS-C with probes the initial (pre-probe) DK levels averaged 13.2% for the first six items and 7.5% for the next five items. Thus, in the no-probe condition DKs rose by a trivial 0.2 percentage points, but under the probe-DKs treatment initial DKs dropped by 5.7 percentage points.

In terms of Don't Knows the 2002 GSS and KNS experiments 1) replicate the 2000 results that KNS produces much higher DKs than the GSS when Don't Know is an explicit response option (i.e. on KNS-B) as opposed to not being verbally or visually offered (as on the GSS), 2) indicate that DK levels vary notably by format, and 3) show that a Web-based, DK format can be devised to replicate the DK levels in an interviewer-administered, in-person survey (KNS-A).

Distributions with Don't Knows Removed

With Don't Knows removed most differences between support for spending levels are small (Table 3). The mean differences is 3.9 percentage points between the GSS and KNS-A and only 6 of the 17 comparisons are statistically significant.

But there are small systematic differences on the items as a whole and moderate-to-large differences in one distinctive sub-set of items. Comparing the GSS to KNS-A (the version with the closest Don't-Know levels to the GSS as well as the largest sample), shows that the GSS tends to get lower proportions saying "too much" on all 17 standard items and higher "too little" proportions on 12 of 17 (see Table 3). However, most of these differences are not

statistically significant.

Additionally, GSS respondents are notably more likely to favor more spending than KNS panelists for big cities (% not too much on GSS - % not too much in KNS = + 6.3 percentage points), drug addiction (+ 7.0 points), Blacks (+ 13.8 points), foreign affairs (+ 8.5 points), and welfare (+ 9.7 points).² These results were similar to the differences found in 2000 when the GSS showed notably greater support for spending for the same five areas (Smith, 2003). Four of these have to do with the problems of the urban underclass and the fifth, foreign aid, with another disadvantaged population.

Several theories have been offered to explain this now replicated difference. Dennis, Li, and Chatt (2004) suggest that the difference results from the difference between the KNS's self-administered mode and the GSS being interviewer administered. They characterize the items as being "potentially controversial and sensitive issues" for which the presence on an interviewer may create a social desirability effect encouraging more sympathetic, pro-spending responses by GSS respondents. The literature (e.g. Tourangeau, Rips, and Rasinski, 2002; Tourangeau and Smith, 1996) shows that self-administered formats typically reduce social desirability effects on the reporting of sensitive behaviors.

²Similar results emerge if the GSS is compared to the average of the KNS results across the four treatments. The GSS finds fewer saying that too much is spent on 16 of 17 items and more saying that too little on 13 of 17 items, the same items show large differences, and the differences are even greater than those based on only KNS-A for all the items with large differences and almost all of them with small differences.

However, these items are not about behaviors and it is not clear that being pro-spending is necessarily the socially desirable response. There is a large anti-spending majority for foreign aid and being for more spending to deal with drug addiction correlates with other items as a punitive rather than a helping and caring response like spending more for Blacks and big cities is. A second theory is that KNS panelists are more conservative on social-welfare issues perhaps because such people are more likely to join and maintain membership in this incented panel. But Dennis, Li, DeShazo, and Cameron (2003) found a slight liberal tendency on the KN panel in an analysis of 8,861 cases. Thirdly, the differences in spending priorities could have to do with differences in survey response rates, conditioning, or other factors. Demographic differences however are not a likely cause since the weighted GSS and KNS profiles on background variables are close (Dennis, Li, and Chatt, 2004).

Questions Wordings

Since 1984 the GSS has had experiments on the wordings of the spending items (Rasinski, 1989; Smith, 1987). Within the GSS the wordings effects are very stable across time and moderate-to-large on some items. As in the 2000 experiments (Smith, 2003), the GSS and KNS showed wordings effects on the same items. But while the direction always agreed, magnitudes often differed. Spending for Assistance to the Poor was more supported than Welfare spending

(GSS +45.9 points; KNS-A +26.4 points). Support for Solving the Problems of Big Cities was greater than Assistance to Large Cities (GSS +27.8 points; KNS-A +11.1). Spending for Dealing with Drug Addiction more than for Drug Rehabilitation (GSS +4.8 points; KNS-A +10.9 points). Thus, while question wordings have a similar directional differences, the magnitude of the effects do differ by mode in several cases.

Conclusion

The 2002 GSS/KNS experiments replicate key findings from the earlier 2000 GSS/KNS research: a) differences in DK levels are confirmed when the Don't Know response option is explicit in the KN mode, b) some spending differences, especially those dealing with the urban underclass and the disadvantaged, are robust, and c) question wording effects are consistent in direction across time and mode, but variant in magnitude by mode. They also demonstrate that the level of DKs are variable by format and that a Web-based format can be devised that produces similar DK levels to those of in-person surveys. Finally, the experiments show that while some measurements are highly comparable across mode (e.g. DK levels when using appropriate formats and most distributions on the spending items), in other cases the results vary notably (e.g. support for spending on the urban underclass and the magnitude of wording effects). Taken together this indicates that differences across modes may tend to be stable, that steps can be taken to reduce or

eliminate some variation (e.g in DK levels), but that much more information is needed on differences in measurement across modes before results from web-based surveys and surveys using other modes can be considered as generally comparable.

Table 1

Don't Know Levels in GSS and KNS

Items	GSS	KNS Treatments			
		A	B	C	D
Spending					
Priorities					
Standard Wordings					
Space	6.4	3.6	21.4	19.9	10.3
Environment	2.2	2.8	10.5	7.0	7.2
Health	1.3	1.5	9.2	6.7	7.7
Cities	8.7	4.8	28.6	24.4	19.8
Crime	2.8	2.4	16.3	11.0	11.3
Drugs	3.1	3.2	18.6	10.0	9.0
Education	0.9	1.2	6.9	2.6	5.1
Blacks	6.8	7.8	28.5	21.1	23.5
Defense	2.3	3.9	13.3	4.1	9.3
Foreign Affairs	3.1	2.4	13.0	5.4	11.4
Welfare	2.9	2.1	12.3	4.4	6.2
Highways	3.2	3.4	16.8	4.5	7.7
Social Security	3.5	2.8	14.0	4.0	7.8
Mass Transportation	5.5	6.3	30.1	11.4	16.4
Parks	2.8	3.0	14.2	3.0	9.7
Child Care	5.8	2.5	19.9	5.0	14.5
Science	6.4	5.3	22.9	5.2	15.5
Variant Wordings					
Space	5.2	3.3	20.7	8.0	12.6
Environment	2.3	2.9	13.2	3.3	9.5
Health	1.3	1.8	9.7	2.1	5.0
Cities	10.5	6.3	30.8	11.3	20.2
Crime	1.8	2.6	15.4	2.8	7.5
Drugs	5.5	3.8	20.3	5.7	11.6
Education	0.9	1.6	6.8	2.2	3.4
Blacks	9.1	8.1	30.1	9.7	20.8
Defense	3.6	3.7	12.9	3.5	11.6
Foreign Affairs	2.7	2.8	12.5	2.4	9.8
Welfare	2.2	2.6	14.3	2.3	6.2
	1364-2762	655	333	334	333

Question Wordings:

We are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or

Table 1 (continued)

about the right amount. First, (READ ITEM A)... are we spending too much, too little, or about the right amount on (ITEM)?

(Standard Wordings are listed first. Variants wording follow the slash. There are only a single wording for items L to Q.)

- A. Space Exploration Program/Space Exploration
- B. Improving and Protecting the Environment/The Environment
- C. Improving and Protecting the Nation's Health/Health
- D. Solving the Problems of the Big Cities/Assistance to Big Cities
- E. Halting the Rising Crime Rate/Law Enforcement
- F. Dealing with Drug Addiction/Drug Rehabilitation
- G. Improving the Nation's Education System/Education
- H. Improving the Condition of Blacks/Assistance to Blacks
- I. The Military, Armaments, and Defense/National Defense
- J. Foreign Aid/Assistance to Other Countries
- K. Welfare/Assistance to the Poor
- L. Highways and Bridges
- M. Social Security
- N. Mass Transportation
- O. Parks and Recreation
- P. Assistance for Childcare
- Q. Supporting Scientific Research

The Ns on the GSS were 1643 for the standard wordings, 1398 for the variant wordings, and 2762 for items L-Q which were asked on both half samples.

KNS Treatments:

A = Instruction on first screen to skip questions to indicate "Don't Know/No Opinion". It read "In this survey we'd like your opinions about some important national issues. There is no right or wrong answer and please the questions based on your knowledge. If you do not know the answer to a question, yo can simply skip the question to indicate 'Don't know' or 'No opinion'." Don't Know option is not presented on screen for individual items.

B = Don't Know presented as response option for each individual item.

C = Don't Know option presented on screen as a response option for each individual item. Then respondents who initially selected Don't Know are prompted to select a response on a follow-up screen that asked, "If you had to choose, are we spending too much money, too little money, or about the right amount of money on..."

D = Instructions on all screens "If you do not know the answer to the question, you can simply skip the question to indicate 'Don't

Table 1 (continued)

Know' or 'No Opinion'." Skipping is done by clicking on a "Next Question" button. Don't Know not listed among response options.

Table 2

Average DK Levels by Mode, Format, and Question Order

Mode/Format	Order		
	11 Standard Items	6 Standard Items	11 Variant Items
KNS-A (Skip-out, First Screen)	3.2	3.9	3.6
KNS-B (Skip-out, Every Screen)	16.2	19.7	17.0
KNS-C (Before Follow-up Probe)	10.6	5.0	4.8
KNS-D (Listed DK Response)	11.0	11.9	10.7
GSS	3.7	4.5	4.1

Table 3

GSS and KNS Spending Priorities with DKs Excluded

Items	GSS			KNS Treatments												
	+	0	-	+	A 0	-	+	B 0	-	+	C 0	-	+	D 0	-	
Spending Priorities Standard Wordings																
Space	11.8	50.4	37.7	11.5	49.8	38.8	15.0	37.6	47.4	10.3	46.3	43.1	11.3	49.5	39.2	
Enviro.	60.0	33.2	6.8	60.5	29.2	10.3	61.7	24.7	13.6	61.6	26.0	12.5	59.6	29.7	10.7	
Health	74.9	21.3	3.9	69.3	25.1	5.6	69.1	23.7	7.2	74.7	20.8	4.6	68.0	27.4	4.6	
Cities	45.4	39.9	14.8	34.4	44.5	21.1	42.9	30.8	26.4	36.9	33.3	29.8	35.4	43.7	20.8	
Crime	57.4	35.8	6.8	51.7	40.8	7.4	54.9	38.0	7.1	51.5	38.4	10.0	55.6	37.7	6.7	
Drugs	59.1	31.2	9.7	45.1	38.2	16.7	47.9	25.9	26.1	49.7	29.4	20.9	45.0	32.3	22.7	
Educ.	73.9	20.7	5.4	73.5	19.9	6.6	69.5	17.6	13.0	73.6	19.4	7.0	66.7	25.6	7.3	
Blacks	32.7	49.0	18.3	23.0	44.9	32.1	24.0	28.1	47.9	26.6	45.1	28.3	21.9	41.6	36.5	
Defense	31.3	46.5	22.3	32.5	44.3	23.1	35.5	45.7	18.8	31.5	48.2	20.3	38.5	42.0	19.5	
ForAffs	6.7	27.8	65.5	6.1	19.9	74.0	4.8	13.9	81.3	6.3	14.7	79.0	8.2	19.7	72.2	
Welfare	21.2	38.2	40.6	18.5	31.2	50.3	13.1	23.9	63.0	15.6	31.8	52.5	14.7	33.1	52.2	
Hghwys.	35.7	51.7	12.7	35.3	51.9	12.8	32.6	50.8	16.6	30.7	54.2	15.1	31.3	55.9	12.7	
SocSec.	60.8	34.6	4.6	62.4	32.3	5.4	66.8	26.1	7.0	61.8	30.1	8.1	58.6	34.5	6.9	
MassTrn	37.0	52.4	10.7	37.5	49.3	13.2	37.2	41.9	20.9	30.0	53.8	16.2	36.9	48.2	14.9	
Parks	35.0	59.5	5.5	37.8	52.8	9.4	30.1	62.7	7.3	37.0	49.1	14.0	30.1	61.4	8.4	
Childcare	59.1	33.2	7.7	53.3	37.2	9.5	53.5	31.4	15.1	48.9	37.4	13.6	48.5	38.3	13.1	
Science	36.4	49.7	13.9	31.6	52.4	16.0	35.7	48.6	15.6	34.2	50.1	15.7	33.7	49.7	16.6	

References

- Bosnjak, M.; Tuten, T.L.; and Bandilla, W., "Participation in Web Surveys: A Typology," ZUMA Nachrichten, 48 (May, 2001), 7-17.
- Couper, Mick P., "Web Surveys: A Review of Issues and Approaches," Public Opinion Quarterly, 64 (2001), 464-494.
- Davis, James A.; Smith, Tom W.; and Marsden, Peter V., General Social Survey, 1972-2002: Cumulative Codebook. Chicago: NORC, 2003.
- Dennis, J. Michael; Li, Rick J.; DeShazo, J.R.; and Cameron, Trudy Ann, "Correcting for Sample Bias in Internet Panel Surveys Based on RDD Sampling," Paper presented to the Joint Statistical Meetings, San Francisco, August, 2003.
- Dennis, J. Michael; Li, Rick; and Chatt, Cindy, "Benchmarking Knowledge Networks' Web-Enabled Panel Survey of Selected GSS Questions Against GSS In-Person Interview," Knowledge Networks Report, February, 2004.
- Dillman, Don A., Mail and Internet Surveys: The Tailed Design Method. New York: John Wiley & Sons, 2000.
- Rasinski, Kenneth A., "The Effect of Question Wording on Public

Support for Government Spending," Public Opinion Quarterly, 53 (1989), 388-396.

Schuman, Howard and Presser, Stanley, Questions and Answers: Experiments on Question Form, Wording, and Context. New York: Academic Press, 1981.

Smith, Tom W., "Educated Don't Knows: An Analysis of the Relationship between Education and Item Nonresponse," Political Methodology, 8 (1982), 47-58.

Smith, Tom W., "An Experimental Comparison of Knowledge Networks and the GSS," International Journal of Public Opinion Research, 15 (2003), 167-179.

Smith, Tom W. "Nonattitudes: A Review and Evaluation," in Surveying Subjective Phenomena, edited by Charles F. Turner and Elizabeth Martin. New York: Russell Sage, 1985.

Smith, Tom W., "That Which We Call Welfare by Any Other Name Would Smell Sweeter," Public Opinion Quarterly, 51 (1987), 75-83.

Smith, Tom W. and Kim, Seokho, "A Review of CAPI-Effects on the 2002 General Social Survey," GSS Methodological Report No. 98. Chicago: NORC, 2003.

Tourangeau, Roger; Rips, Lance; and Rasinski, Kenneth A., The Psychology of Survey Response. New York: Cambridge University Press, 2000.

Tourangeau, Roger and Smith, Tom W., "Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Content," Public Opinion Quarterly, 60 (1996), 275-304.