Context Effects Resulting from the Revision of the GSS Core in 1994

Tom W. Smith

National Opinion Research Center
University of Chicago

January, 1995

GSS Methodological Report No. 83

This research was done for the General Social Survey project directed by James A. Davis and Tom W. Smith. The project is supported by the National Science Foundation, Grant No. SES-9122462.

Introduction

NORC's National Data Program for the Social Sciences has been monitoring change since 1972 with its General Social Surveys (GSSs). The GSSs are cross-sectional, in-person surveys of adults living in households in the United States. The GSSs have been conducted annually from 1972 to 1994 with the exceptions of 1979, 1981, and 1992. For full details see Davis and Smith, 1994.

The GSS project has striven to faithfully measure true change by the strict replication of measurement procedures such as sampling, question wording, item order, coding protocols, and so forth. When necessary changes in procedures have been implemented, either experimental designs have been carried out to assess any measurement effect or other research has been conducted to ascertain the impact of measurement variation on the time series. Past examples include 1) the split-ballot experiments in 1975 and 1976 when the GSS switched from probability with quota sampling to full-probability samples (Smith, 1979; Stephenson, 1979), 2) the split-sample frame experiments in 1983 and 1993 when NORC revised its sample frame according to the latest Census (Smith, 1990), 3) various studies of question wording (Smith, 1984 and 1987; Rasinski, 1989), 4) examinations of the impact of question layout and format (Smith, 1993), and 5) experiments on questionnaire context and order (Smith, 1983; 1984; 1986; 1991a; 1991b).

Of particular relevance to the present study is the work carried out when the GSS shifted in 1988 from rotating items across years to rotating items across ballots within a year (Smith, 1988). This design shift allowed the experimental assessment of context effects across all GSS items (Smith, 1991a) and how these effects may have affected time series (Smith, 1990).

In 1994 another major shift in design was implemented when the existing GSS core of 60 minutes was reduced to 45 minutes ("List," 1993). The deletion of a little over one-quarter of the existing core, plus the addition of a small number of new core items meant that the context in which retained items appeared was notably altered. Three steps were taken to minimize the distortion to the continuing time series of these changes in context: 1) items were not deleted from the middle of explicit scales, either entire scales were dropped or, in a few cases, only items at the end of scales were removed, 2) new items were either placed at the end of sections covering the same subject matter or the end of the core as a whole, and 3) an experimental design was instituted to assess context effects from the core reduction.

1994 Context Experiment

The under the context experiment one half of sample A (approximately 750 cases) received the old core and the other half got the new core (750 cases). In addition, all sample B cases (approximately 1,500) also were administered the new core. This design permits comparison of responses to items remaining in the core (and thus in both the old and new core) under both their

original (old core) context and their revised (new core) context. Of course since most GSS items appear on only two of three ballots, this means that in most cases there are approximately 500 cases rather than 750 under the old-core condition and 500 to 1500 cases under the new-core context (depending on whether only sample A or samples A and B are utilized).

Analysis

Analysis proceeded in three steps. First, retained, old core items that appeared in the main questionnaires (e.g. excluding items from the sample frame and household enumeration form that were not affected by the change in core content) and other items similarly affected by the core changes were cross-tabulated by context. Second, for those showing statistically significant variation, the proximate differences in context were delineated and their probable effects assessed. Finally, to examine which statistically significant differences probably represented real effects rather than chance variation, a) further analysis was done within ballots and sub-groups and b) the consistency of context effect across items that had dealt with related topics, were correlated with one another, and had similar locations were examined.

Overall of 192 contextual comparisons, 14 or 7.3% showed statistically significant differences at the .05 level (Table 1). Since 9-10 differences would have been expected by chance, this suggests that 4-5 differences are probably the result of changes in context.

Of the 14 variables showing statistically significant differences across context, seven were deemed to have no notable variation in the immediately preceding context, one had different, immediate contexts, but the variation in context did not seem clearly connected to the differences in responses, and six had context variations that could plausibly be related to observed differences in context (Table 2).[1]

For 9 of the 14 items no related items had context effects. For example, while extra-marital sex was less disapproved of under the new core, there were no significant differences for approval of teenage, premarital, or homosexual intercourse and two of these related three items showed insignificant differences in the opposite direction. Two items had no related items in close proximity (GETAHEAD and RICHWORK), one item (SPKCOM) showed some indication of related effects, and two items (HEALTH and LIFE) had consistent effects with each other.

---

[1]Assessments of the relevancy of contexts are based on informed judgments based on a review of the extant literature on context effects and past GSS experiments (Smith, 1991a; 1991b). Strictly defined context varies for all items. For a number of the items however, no variation in context occurs during 5-15 minutes of interviewing immediately prior to the occurrence of the item.

We eliminated the seven items with no meaningful variation in context and showing no consistent context effects among related items as unlikely to denote real context effects and examined the remaining seven items which showed possibly meaningful differences in context and/or similar effects on related items.[2]

First, frequency of church attendance (ATTEND) and of personal prayer (PRAY) are both about the frequency of religious behaviors and occur within five items of each other. However, since the observed effects point in opposite directions (more religious on the old core for prayer and less for church attendance), they tend to contradict rather than support one another. Since the context does not vary for 11 items prior to the church attendance item (and then varies only by whether two or six items on pornography appeared), this context effect is considered as of low likelihood. For prayer, however, the immediately preceding item varies. Images of God (i.e. whether God is a mother-father, judge-lover, friend-king, or master-spouse) appears on the old core and on the new core is replaced by two world views items (the world is good or evil and human nature is good or corrupt). The more frequent mention of praying under the old-core, images-of-God condition may result from the greater, prior focus on God, the recipient of one's prayers.

Second, personal evaluations of one's state of health (HEALTH) and the activity in one's life (LIFE) both deal with self-evaluation and on Ballot A appear within two items of each other. Thus for Ballot A the observed context effect is consistent in direction (more positive on the new core than the old) and in the same approximate location, so the results tend in a general sense to confirm each other. For Ballot A a five-item scale rating job values appears immediately before the life evaluations (and thus four items before health) on the old core, while this item is omitted on the new core. Also, seven items before the life item a three-item anomia scale appears in the old core, while three new race relation items appear on the new core. On Ballot C the life item is also influenced by the immediate prior appearance of the five job-value items, but the anomia items are not nearby for either the old or new core. It is plausible that the focus on job values may lead to life being judged as less exciting and on Ballot A, which shows the larger effect, the anomia items may also depress choices of the exciting response. The connection to health is less obvious, but the same factors than lead to life being rated as less exciting may reduce expressions that one's health is excellent.[3]

Third, on Ballot C the questions immediately prior to the item

---

[2]These items were not separated on the basis of their probability levels. On average the seven dismissed items were significant at the .025 level and the seven examined items at the .021 level.

[3]In addition, measures of psychological well-being and of personal evaluation have shown a sensitivity to context effects in general (Smith, 1986, 1991a, 1991b).

on whether one gets ahead in life primarily through hard work or luck differs little and there is no statistically significant difference across the two cores. On Ballot A however getting ahead is preceded on the old core by three anomia items and then a large race battery. For the new core however the three anomia items are replaced by three additional race items including one that asks whether one agrees or disagrees that "Irish, Italian, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without special favors." Selections that one gets ahead through hard work are 16.2 percentage points higher on the new core (prob.=.0002) than on the old core. It is likely that the assertion of the work ethnic in the new race item, either activated images of hard work or otherwise enhanced the judgment that one advances via hard work.[4]

Fourth, a desire to still work if one became rich (RICHWORK) does not show any immediate variation in context on Ballot C, but does show some differences on Ballot A. Two items about the chances of losing one's job and finding as good of a job precede the work if rich item on both the old and new core and then occupational and work force participation questions appear. On the new core however two extra items about weeks worked and level of employment last year appear between the job loss and occupational items. However, since the context effect is larger for the context showing the no proximate difference (-9.6 percentage points, prob.=.054) than for the more variant context (-4.4 percentage points, prob.=.399), the impact of the added items seems unlikely.

Finally, on both Ballots A and C allowing free speech for Communists (SPKCOM) is preceded by three civil liberty items on racists and three on those opposed to organized religion. Then in the old core comes an item on evaluating whether Communism is a good or bad form of government, while this item is omitted on the new core. (The next preceding group of items do not differ between the old and new core.) Obviously prior judgments about how good a form of government Communism is might influence later evaluations of free speech for Communists. There is mixed evidence for this possible connection. None of the other 14 civil liberty items, including items on letting a Communist teach in a college and have a book in the public library, show any significant differences. However, the insignificant differences consistently show slightly greater tolerance in the new-core context. But even when scales of the five free-speech items or the three Communism items are constructed, there are no significant differences by context (respectively prob.=.114 and .161).

Overall, there is strong evidence that a context effect occurs on getting ahead through hard work (although interestingly only on one Ballot). There is some reason to believe that frequency of

---

[4]It is unlikely that the three anomia items on the old core reduced mentions of hard work because mentions on Ballot A on the old core (with anomia) are not significantly lower than on Ballot B on the old core (without anomia).

praying, health status, life assessment, and free speech are experiencing real context effects. Finally, it appears unlikely that the differences on church attendance and continuing to work result from the differences in context. Thus, we find evidence of from one to five context effects which is consistent with the four to five differences predicted by the overall distribution of differences.

Conclusion

Fortunately only a few time series appear to be disrupted by context effect.[5] The results are consistent with past research that as long as explicit scales are not disrupted context effects occur once out of every 40-60 cases (Smith, 1991a). For the five probable or possible context effects researchers will have to be judicious in their time-series analysis and consider whether an adjustment based on the 1994 context experiment might be best used when studying trends.

[5]Of course there may well be many relatively minor effects, effects too small to reliably detect given the precision provided by the size of the samples employed in our experiment.

Table 1

Statistically Significant Differences by Context

Variables	Category	Old Core	New Core	Prob.[a]
ATTEND	% never attend church	17.2	13.8	.032
PRAY	% pray daily	61.8	54.5	.015
HEALTH	% excellent health	28.0	35.4	.013
LIFE	% life is exciting	43.6	53.0	.004
RICHWORK	% continue to work	69.3	61.9	.042
GETAHEAD	% get ahead by hard work	65.6	75.1	.003
CONLABOR	% grt deal of confidence	11.7	10.4	.010
XMARSEX	% extra-marital sex always wrong	81.7	79.1	.010
COOP	% interviewer rated as friendly/cooperative	77.6	74.7	.048
MARFIN	% agree, marriage gives financial security	20.5	14.5	.045
MARNOMAR	% agree, bad marriage better than no marriage	4.3	1.5	.005
RWRKBABY	% worked when child was pre-schooler	59.8	59.1	.040
ACQNTSEX	% had sex last year with acquaintance (if partners besides spouse/regular partner)	18.9	37.0	.018
SPKCOM	% allowing Communist to give speech	65.3	71.3	.041

[a]Probability tests were based on sample A cases with Ns generally between 1,000 and 1,500. All categories, not just the category displayed above, were used in the calculations of statistical significance.

Table 2

Summary of Judgments about Items Showing Significant Differences

Variables	Categories	Context Different	Related Items Show Effect
ATTEND	% never attend church	M	N
PRAY	% pray daily	Y	N
HEALTH	% excellent health	Y	Y
LIFE	% life is exciting	Y	Y
RICHWORK	% continue to work	Y	X
GETAHEAD	% get ahead by hard work	Y	X
CONLABOR	% grt deal of confidence	N	N
XMARSEX	% extra-marital sex always wrong	N	N
COOP	% interviewer rated as friendly/cooperative	N	N
MARFIN	% agree, marriage gives financial security	N	N
MARNOMAR	% agree, bad marriage better than no marriage	N	N
RWRKBABY	% worked when child was pre-schooler	N	N
ACQNTSEX	% had sex last year with acquaintance (if partners besides spouse/regular partner)	N	N
SPKCOM	% allowing Communist to give speech	Y	M

Y=yes; M=maybe; N=No; X=not applicable - no related items in close proximity

References

Davis, James A. and Smith, Tom W., General Social Surveys, 1972-1994: Cumulative Codebook. Chicago: NORC, 1994.

"List of Deletions from GSS Core," GSSNews, No. 7 (September, 1993), 1-5.

Rasinski, Kenneth A., "The Effect of Question Wording on Public Support for Government Spending," Public Opinion Quarterly, 53 (Fall, 1989), 388-394.

Smith, Tom W., "Children and Abortion: An Experiment in Question Order," GSS Methodological Report No. 27. Chicago: NORC, 1983.

Smith, Tom W., "Context Effects on the General Social Survey," in Measurement Errors in Surveys, edited by Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman. New York: John Wiley & Sons, 1991a.

Smith, Tom W., "Little Things Matter: A Sampler of How Differences in Questionnaire format Can Affect Survey Responses," GSS Methodological Report No. 78. Chicago: NORC, 1993.

Smith, Tom W., "A Preliminary Analysis of Methodological Experiments on the 1984 GSS," GSS Methodological Report No. 30. Chicago: NORC, 1984.

Smith, Tom W., "Rotation Designs on the GSS," GSS Methodological Report No. 52. Chicago: NORC, 1988.

Smith, Tom W., "Sex and the GSS: Nonresponse Differences," GSS Methodological Report No. 9. Chicago: NORC, 1979.

Smith, Tom W., "Some Thoughts on the Nature of Context Effects," in Context Effects in Social and Psychological Research, edited by Norbert Schwarz and Seymour Sudman. New York: Springer Verlag, 1991b.

Smith, Tom W., "That Which We Call Welfare by Any Other Name Would Smell Sweeter: An Analysis of the Impact of Question Wording on Response Patterns," Public Opinion Quarterly, 51 (Spring, 1987), 75-83.

Smith, Tom W., "Timely Artifacts: A Review of Measurement Variation in 1972-1989 GSS," GSS Methodological Report No. 56. Chicago: NORC, 1990.

Smith, Tom W., "Unhappiness on the 1985 GSS: Confounding Change and Context," GSS Methodological Report No. 34. Chicago: NORC, 1986.

Stephenson, C. Bruce, "Probability Sampling with Quotas," Public Opinion Quarterly, 43 (Winter, 1979), 477-496.