Context Effects in the General Social Survey

Tom W. Smith

NORC

University of Chicago

August, 1988
Revised November, 1990

Paper presented to the International Conference on Measurement Errors in Surveys, November, 1990, Tucson.

Publication Notes: A revised version of this paper was published 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, 1991.

This research was done for the General Social Survey project directed by James A. Davis and Tom W. Smith. It revises and expands upon GSS Methodological Report No. 55. The project is funded by the National Science Foundation Grant SES-8747227.

6.1 INTRODUCTION

After a half century of study, question order is probably the least developed and most problematic aspect of survey research. As Schuman and Presser remark in their work on survey methodology (1981, p. 77), "Overall, order effects...constitute one of the most important areas of methodological research. They can be very large [and] are difficult to predict." This perplexity is shared by Bradburn (1985, p.302) who observes, "No topic in questionnaire construction is more vexing or resistant to easy generalization than that of question order" and by Groves (1989, p. 479) who notes, "(T)here seems to be no general theory that predicts when such effects are to be expected and when they should not be expected."

In part because of our limited ability to predict their occurrence, there is some disagreement in the literature over how common context effects. Tourangeau, et al. conclude that "The literature on survey context effects may create the impression that such effects are relatively rare, involving items on a few

scattered issue. These results here indicate otherwise... (Tourangeau, Rasinski, Bradburn, and D'Andrade, 1988, pp. 22-23)." This impression of pervasiveness is supported by numerous instances in which changes in question order have upset time series or otherwise caused other undesired measurement variations (Smith, 1986; Smith, 1988c; Astin et al., 1988; Cowan, Murphy, and Weiner, 1978; Roper, 1990; Johnson and Bachman, 1980; Turner and Martin, 1984; and Gibson, Shapiro, Murphy, and Stanko, 1978).

Schuman and Presser, on the other hand, reach a conclusion that at least differs in emphasis - "Question order-effects are evidently not pervasive...but there are enough instances to show they are not rare either (1981, p. 74)." This non-pervasive impression is supported by numerous failures to produce context effects in experiments designed to do so (Schuman and Presser, 1981; Smith, 1983a; Turner and Martin, 1984) and by the ability of different houses to produce similar marginals when the same questions, but different question content (as well as other variations), existed (Turner and Martin, 1984; Smith, 1978; Smith, 1982).

To date only one study has conducted a general search for context effects. Schuman and Presser (1981) examined the 1971 Detroit Area Study (DAS). The DAS used split ballots in order to accommodate various experiments in either question order or wording. They looked at 113 attitude items that were not the designed objects of these experiments, but appeared after the experiments and thus varied in context due to the prior experiments. Apparently using simple random sample (SRS) assumptions, they found eight significant differences at the .05 level, just two above what chance would predict. Their inspection of these eight suggested that three probably represented real effects and the rest were due to sample variation.

In this paper we conduct a similar analysis using the 1988 and 1989 General Social Surveys (GSS).

6.2 THE GSS DESIGN

The GSS has monitored trends in America since 1972. The GSS has striven to measure true change undistorted by measurement artifacts (Smith, 1986; 1988c). To avoid distortion in the time series, the project has tried to keep all facets of data collection and processing constant; using identical sampling procedures, question wordings, data processing, and so forth. When necessary changes have occurred, such as the periodic updating of the sample frame, the GSS has utilized calibration designs or other features that allow the separation of measurement variation from true change and the maintenance of a consistent time series.

One of the measurement variations that has been most difficult to control for has been order or context (Smith, 1986). The exact replication of question order has not been possible because of several other features of the GSS:

1) Until 1988 the GSS employed a rotation design under which most GSS items appeared on two out of every three surveys. (For details of the rotation design see Smith, 1988b.) This meant that the same items appear only on every fourth survey and, as a result, it was impossible to exactly duplicate order across adjoining surveys.

2) The content of the survey has changed because of the addition and deletion of items. Such changes were especially large between 1972 and 1973 when the survey length was substantially increased and between 1984 and 1985 when items were dropped or switched from permanent status to rotating in order to open up space for annual supplements.

3) Methodological, split-sample experiments on question wordings, other features, and context itself have altered the order of questions.

To minimize the unavoidable variation in order caused by these factors, the GSS has tried to 1) keep order as constant as possible within the rotation patterns so that the order of the first and fourth, second and fifth, and third and sixth years was the same, 2) maintain the order of all scale items unaltered (e.g. the seven-item abortion scale or the 15-item national spending scale), 3) place new items either at the end of the survey (as in the case of the topical and cross-national supplements) or at the end of individual scales (as when an additional abortion question ,ABANY, was added to the abortion scale), 4) consider the possibility of order effects when constructing the questionnaire each year, 5) check all "blips" in the time series for possible context effects and 6) study context effects via split-sample experiments both to extend our general knowledge of context effects and to test for specific effects and possible distortions to the time series suspected on the GSS (Smith, 1986; 1984; 1983b; 1983a; 1982a; 1981).

With the switch from the rotation across time design utilized by the GSS until 1988 to the split-ballot design employed since 1988, the GSS is able to test for possible context effects by comparing response distributions across ballots. (For details of the new design see Smith,1988b.) In brief, under the new split-ballot design items that would have appeared on two of the next three surveys instead appear on two of three random sub-samples (called ballots), each with one-third of the cases in a particular year. Thus items that would have appeared in 1988 and 1989 and would have been off-rotation in 1990 instead appear in 1988, 1989, and 1990 on ballots A and B. Items that would have appeared in 1988 and 1990 and would have been off-rotation in 1989 instead appear on all three years on ballots A and C. Likewise items off in 1988 and slated for the 1989 and 1990 surveys appear on all three surveys on ballots B and C.

Retrospectively the switch means that ballot A in 1988/89 represents the 1982 and 1985 surveys, ballot B represents 1983 and 1986, and ballot C represents 1984 and 1987. (Ballots also resemble

earlier years in the GSS series, but the closeness of the match decays over time due to the addition and deletion of items and other variations. The matches between the 1982 through 1987 surveys and the ballots in 1988 and 1989 are much closer than the match with earlier years.)

This switch from a rotation to a split-ballot design makes it possible to experimentally test for context effects by comparing items across ballots. Items that followed the previous rotation design will appear on two ballots (AB, AC, or BC), while permanent, non-rotating items will appear on all ballots (ABC). Simple comparison of the distributions across ballots allows for the detection of context effects. And since the ballots largely duplicate the order of questions in previous years, one can usually generalize these findings to changes across past GSSs. In effect, differences across ballots on the 1988 and 1989 surveys should duplicate changes across rotation years in past GSSs. Consequently, order effects on the 1988/89 GSS identify order effects that have been distorting the time series on past GSSs.

6.3 ANALYSIS OF CONTEXT EFFECTS

Table 6.1 shows the results of crosstabulating the 502 variables on the 1988 GSS and the 442 variables in 1989 with ballot. Overall using SRS assumptions, 9.0% of the comparisons in 1988 and 8.1% of comparisons in 1989 are significant at the .05 level. Using a design effects adjustment of .667 lowers the percentage of significant differences to only 2.8% in 1988 and 3.2% in 1989. Table 6.1 breaks the GSS variables into five categories. Two of the categories, variables from the sample frame and household enumeration form (HEF) and variables occurring at the beginning of each ballot, do not vary in context across ballots. Two other groups, variables from the topic modules and variables from the sexual behavior supplement are both internally consistent in order across ballots and separated from the rest of the questionnaire by a buffer of identical questions. (In 1988 the religious belief and behavior module was preceded on all ballots by four questions on attitudes towards science, questions on past and present religious preference, a question on strength of religious affiliation, and a series of questions about changes in religious preference in the past. In the 1989 forbid/allow experiments followed a number of identical demographics on both ballots. In each year the sexual behavior supplement is preceded by the 15-minute long International Social Survey Program (ISSP) supplement. In 1988 the ISSP module was on the impact on the family of the changing labor force participation of women and in 1989 it was on orientations towards work.) For these four groups the likelihood of context effects is nil to negligible. The fifth group consists of all other variables which to a greater or lesser extent differ in their order across surveys.

The aggregate level of context effects is very similar in 1988 and 1989. In both years there are very few significant differences among the context-similar groups while more significant differences appeared among the context-different variables. For 1988 the SRS

results in Table 6.1 find almost as many significant differences for the context-similar categories (8.3%) as for the context-different categories (9.2%), but the adjusted significance tests show that most of the context-similar variables had differences of only borderline significance. With adjustments only one of the context-similar variables (0.7%) showed a statistically significant difference with 13 of the context-different variables (3.6%) still registered as significantly different. For 1989 both the SRS and adjusted probability levels show few (1.6%) significant differences among the context-similar questions. Among the context-different questions 9.2% showed significant differences prior to adjustments, but only 3.4% were significant after adjustment.

6.4 ASSESSING SUBGROUP DIFFERENCES

Since fewer than 5% of the comparisons in both 1988 and 1989 were significant in the adjusted comparisons, it might be possible to argue that no true differences are occurring, that all differences are occurring by chance. Rather than make this general probabilistic conclusion, we evaluated each of the 28 statistically significant differences by a) comparing the 1988 and 1989 results both to one another and to the 1982-1987 surveys, b) evaluating the differences in context across the ballots, and c) searching for conditional or associational effects consistent with our proposed explanations for context effects.

Each of these evaluations involves some difficulties. The comparison of 1988 and 1989 is generally straight forward since most variables appeared both years and context was usually identical across years. The pre-1988 comparisons are more problematic however. We carried out the pre-1988 comparisons by grouping together the two most recent years representing the same rotation of ballot (82 and 85; 83 and 86; 84 and 87). We used only the two most recent rotation cycles because changes from the addition and deletion of items and other reasons make earlier years less comparable to the 1988 and 1989 ballots than more recent years. Since the pre-1988 comparisons are usually based on nearly 3,000 in each "ballot" condition (compared to about 500 on each ballot in 1988 or 1989), rather small differences can be statistically significant. More importantly, collapsing years into quasi-ballots does not eliminate time as a factor. Not only would items varying across time in happenstance with rotation create associations, but items showing secular trends also relate to rotation. (The collapsing of 1982, 85; 1983, 86; and 1984, 87 reduces somewhat the linear association between a variable and time, but each rotation still averages a year latter in time and thus rotation year and time are still associated.) In addition, as noted above, the context in 1982-1987 is not exactly the same as in 1988 and 1989 due to the addition/deletion of items, experiments, and related reasons.

Assessing the cause of context effects is also difficult because such effects are complex and imperfectly understood (Schuman and Presser, 1981; Smith, 1989; Tourangeau, 1987; Tourangeau, Rasinski, Bradburn, and D'Andrade, 1988; Tourangeau and

Rasinski, 1986). While context effects usually occur between related items, the interconnection between questions can be subtle and hard to predict. In addition, for most of the context-different comparisons in Table 6.1 (and in particular the 28 significant differences) preceding context is complex, involving many questions covering numerous subjects. Although the question(s) inducing a context effect usually closely precede the affected question, this does not have to be the case. (See Bishop, Oldendick, and Tuchfarber, 1982; Schuman, Kalton, and Ludwig, 1983 for examples of long-distant order effects.) It is effectively impossible to assess in detail the entire context in which each of the affected questions appeared. We examined and noted the four immediately preceding questions (which because of sub-questions and indices often involved many variables) and generally searched an additional 15-20 questions for possible context inducing questions.

Finally, to check on the plausibility of these context explanations, we examined if any other association could be established between the hypothesized triggering variable(s) and the affected question. Context effects come in various forms and in many instances there is no detectable connection between the two variables other than the marginal changes on the affected variable (Smith, 1982). There are two types of context effects however that not only lead to a change in the marginal distribution of the affected variable, but also have other connections between the variables. First, in the case of associational order effects, the relationship between the variables is increased when the context effect is operating. This can occur when the preceding variable creates a frame of reference by which the subsequent variable is evaluated (Smith, 1981). Second, with conditional order effects the impact on the subsequent question occurs only among people giving certain responses to the prior question or, in the extreme case, the context effect is in opposite directions depending on how the preceding question was answered (Smith, 1982; 1989).

We first considered whether the explanation for the context effects implied conditional and/or associational effects. In the majority of cases, the proposed explanations did not suggest such effects. In some other cases conditional or other effects were implied by our explanation, but the triggering variable appeared on only one ballot so we could not empirically evaluate our explanation. When a conditional or associational effect was predicted and it was possible to test it, we carried out the appropriate comparisons on both the pre-1988 and 1988-89 pooled data. For example, we examined whether helpfulness was more associated with violence questions when these questions were in close proximity (an associational effect) and whether support for home care of elderly parents was higher after the helpfulness item only among those expressing the belief that people are helpful (a conditional order effect).

Table 6.2 shows how the variables differed across ballots in 1988 and 1989 and across rotation years from 1982 to 1987. Table 6.3 summarizes the data from Table 6.2 and makes a judgment on what differences represent real order effects.

Table 6.3 shows the 31 variables that statistically varied in 1988 or 1989 or were candidates from the pre-1988 analysis. (6 from pre-1988, 14 from 1988, 14 from 1989 minus three that were on both the 1988 and 1989 lists.) Leaving aside the two scales which overlap with individual questions, there are 29 questions. Inspection of the pre-1988, 1988, 1989, and pooled 1988-89 differences across ballots suggests that 11 or 12 (Yes or Yes?) represent true context effects, 2 are uncertain, and 15 (No or No?) probably represent chance occurrences.

For 13 of the 15 variables judged not to represent true context effects, the pooled 1988-89 differences are not statistically significant (when adjusted) and the direction of the pre-1988, 1988, and 1989 differences are not consistent. In addition, context is fairly similar in each case and no probable source of an order effect was detected. For two variables, trying to convert someone to Christ (SAVESOUL) in 1988 and allowing a racist to speak in public (ALLOWRAC) in 1989, the variables appeared on only one survey and therefore no cross survey check is possible. They are the only two context-similar variables to show significant differences and neither seems a very likely candidate for a true context effect. SAVESOUL has identical context on all three ballots in 1988 for at least the previous 28 questions and no variables closely related to it in either placement or association show signs of a context effect. Similarly, for ALLOWRAC in 1989 the immediate context is similar on both ballots and none of the five related variables in the forbid-allow experiment show any context effects. In addition, examination of the forbid-allow experiments for possible context effects from prior Stouffer questions revealed no signs of a context effect.

For two variables, we have considered the context effects as uncertain. For the first, membership in nationality groups (MEMNAT), the pattern is really similar to several that we judged not to represent real effects. We have placed these variables in the uncertain category primarily because we found another membership variable probably has a real context effect and therefore are accepting the idea that something about the different contexts influences at least one membership item and therefore might be influencing other membership variables as well. The second uncertain variable, the Stouffer item on letting a militarist have a book in the public library (LIBMIL), is of borderline significance for the pooled 1988-89 sample and the differences are always in the same directions. On the other hand, the immediate context is very similar across ballots.

6.5 SUSPECTED CONTEXT EFFECTS

That leaves 11 to 12 variables (11 Yes's and 1 Yes?) that are significantly different for 1988-89, in the same direction in the pre-1988, 1988, and 1989 comparisons, and have plausible, if unproven, context explanations for the differences and/or a past history of being prone to context or other method effects. Not all of these effects are equally probable or of similar magnitude, but

based on the criteria listed above they are considered as probably representing real context effects.

First, belief in the helpfulness of people (HELPFUL) is higher when following items about leisure activities and lower when following items about economic standing and violence. We believe that the economic items might focus thoughts on non-helpful behavior (competition and self-interest) and/or the violence approval questions might emphasize criminal and aggressive behavior and thus reduce one's evaluation of people's helpfulness.

Second, trust in people (TRUST) is less when following three crime control variables than when following a mixture of questions on how exciting one's life is, job values, reasons for advancement, and homosexuality. As with helpfulness, we believe that the crime items might have reduced evaluations of the trustworthiness of people.

Third, there is lower satisfaction with the local community (SATCITY) when questions about the helpfulness and fairness of people and leisure time activities come first than when 15 Stouffer civil liberties questions, a question on Communism, an item on equalizing wealth, and a 13-part question on confidence in institutions come first. Past research has identified happiness and satisfaction items as susceptible to order effects (Smith, 1986; McClendon and O'Brien, 1983; Schuman and Presser, 1981). In this situation however the source of the impact is not readily apparent. One possibility is that a contrast effect after the Stouffer items raised positive evaluation of one's community (i.e. in contrast to the various politically or socially deviant groups asked about in the Stouffer questions, aspects of one's immediate life may be viewed more positively). We believe that the effect is greater for satisfaction with community rather than the other satisfaction items because satisfaction with community is the first satisfaction item. (For a similar first only effect see Smith, 1981).

Fourth, satisfaction with ones friends (SATFRND) is also apparently influenced by the same factors mentioned in the discussion of SATCITY above.

Although the other three satisfactions items (satisfaction with family, leisure time, and health - SATFAM, SATHOBBY, SATHEALT) do not show clear, individual context effects, it appears that all are being influenced in a similar way as these variables. For the five satisfaction items in the pre-1988, 1988, and 1989 comparisons, satisfaction is always greater when following the Stouffer questions than in the other context.

Fifth, the anomia item on it hardly being fair to bring a child into the world (ANOMIA6) is more agreed with when following the spending priority items than when coming after race relation items. (Another anomia item immediately preceded ANOMIA6 on both versions.) The no children anomia item has been shown to be highly sensitive to context effects in the past (Smith, 1983a), but we are unsure what context in the current situation depressed or elevated

anomia.

Sixth, the anomia item that states public officials do not care about citizens (ANOMIA7) is more agreed with when following the spending priority items than when coming after race relations. Since ANOMIA7 immediately followed ANOMIA6 we assume that a common context is influencing both.

Seventh, confidence in the people running Banks and Financial Institutions (CONFINAN) is higher when the preceding context includes questions about number of children, unemployment history, work hierarchy, and occupational questions than a version that asks about social class, personal financial status, union membership, and occupation. (On both ballots 12 other confidence items appear before the financial item.) We speculate that confidence in Banks and Financial Institutions may be lowered by the prior questions on personal financial conditions. This explanation is supported by the fact that when asked first about personal finances there is a positive correlation between personal difficulties and low institutional confidence, but when the personal financial questions did not appear first, personal financial evaluations and institutional confidence are not related. For example, when the personal financial questions came first, among those reporting improvements in their financial situation 30% had a great deal of confidence in financial institutions. Among those reporting no change in their personal condition 25.5% had a great deal of confidence and among those reporting being worse off only 20% had a great deal of confidence. When the personal financial question did not appear first, there was no association (% great deal of confidence if better (18%), same (22%), and worse off (21%)).

Eighth, confidence in the people running Organized Religion (CONCLERG) is also higher in the latter condition described above. While we believe that there is a context effect affecting confidence ratings in general (see below), we do not know why significant individual effects showed only for these two items. They are not closely related institutions and appear at nearly opposite ends of the lists of institutions. In addition, confidence in religion does not show the conditional association to ones personal financial situation that the financial item does.

It appears however that context is impacting on the confidence scale as a whole rather than just on the two variables that show clear, individual differences. An additive confidence scale has marginally significant context differences for 1988-89 and shows the same difference in the pre-1988 data. Across the 13 confidence items in the pre-1988, 1988, and 1989 comparisons confidence is greater in the version that omits the subjective financial questions 37 times out of 39. In addition, the confidence questions have been shown to be susceptible to other order effects (Smith, 1981).

Ninth, support for having elderly parents live with their children (AGED) is less after eight feminism questions (and items on school prayer and the UN) than after items on helpfulness and

economic standing. We hypothesize that either exposure to feminism items emphasizing modern roles reduced support for traditional, family care or that an expression of the helpfulness of people increased willing to accommodate elderly parents (or both may be acting).

Tenth, more memberships in Church-affiliated Organizations (MEMCHURH) are reported after a battery of questions on pornography than after questions on smoking, drinking, and education. Since religious membership is preceded on both ballots by membership questions on 14 other groups, the impact apparently occurs across this buffer. Once again we are uncertain of why the context occurs. It is known that the church question is ambiguous and subject to varying interpretations and thus may be subject to contexts that activate these various understandings (Smith, 1988a). There might be a self-presentation effect of mentioning more religious affiliations after the pornography questions. Or the "sin" questions may lead people to define Church-affiliated Organizations in a wider fashion thinking perhaps of their religious affiliation in general and/or perhaps drawing in morality-oriented groups in addition to religions.

Eleventh, more attendance of a racially integrated church (RACCHURH) is reported following questions about outlawing interracial marriages and support for open housing than about degree of neighborhood integration and approval of school integration. One possible explanation is that the neighborhood question frames the church integration question such that people tend to think of their local congregation ("church"), while in the other context more people think of their denomination ("church"). Since a local congregation is less likely to be integrated than a denomination, more integration would be reported in the former case than the latter.

Finally, selecting a "feeling of accomplishment" as an important preference in a job (JOBMEANS) is more common following questions on sexual and reproductive issues than when after anomia items and racial items. (In both cases the job value question was immediately preceded by a question on whether getting ahead in life depends on luck or hard work.) This item is of only borderline plausibility. The effect may emerge from the negatively phrased anomia items reducing mentions of the somewhat idealistic "accomplishment" preference and promoting the selection of more pragmatic selections such as high income, job security, and a chance for advancement. However, the effect is not consistently greater among those with anomic feelings which questions this explanation.

6.6 CONCLUSION

Overall the number of context effects created by the rotation design employed by the GSS prior to 1988 are minuscule; 11-12 probable order effects out of over 500 variables. Even this dirty dozen was concentrated among a few topics: anomia, misanthropy,

satisfaction, and institutional confidence accounting for 8 of the 12 context effects. Distortions to the existing time series are also generally small in magnitude. Of the 12 probably real effects the average size is 7.5 percentage points (using pooled 1988-89 figures) with the largest being trusting people (TRUST) at 11.0 percentage points and the smallest being the less certain job values question (JOBMEANS) at 2.6 percentage points.

Together with the Schuman and Presser DAS study, our GSS study suggests that unanticipated context effects might occur once out of every 40-60 questions. In one regard however this is probably an underestimate, since on the GSS, and presumably on the DAS, batteries of questions on one topic (e.g. the seven items on abortion or the 13-item confidence battery) were asked in a block and not varied across ballots. Since context effects are most likely to occur between closely related items, the failure to experimentally vary items within topical blocks, probably underestimates the frequency of context effects.

In addition to the commonness of context effects, the review of context effects on the GSS due to ballot position suggests certain characteristics of questions that make them prone to such effects.

First, many of the affected items seem to be rather general in their scope. For example, the items on helpfulness and trustworthiness ask for global judgments on the nature of man and one of the anomia item asks people to consider "the way things look for the future." All of these items theoretically make relevant nearly a lifetime of memories of the most varied and complex sort. Since the mind is neither disposed to such massive processing of memory nor to a random sampling of same, it is likely that the sampling of memories would be notably influenced by those memories and thoughts triggered by prior questions and that such priming would affect one's subsequent response to such general questions. While context effects do not always involve only such memory-overloading questions (e.g. the Russian/American reporters items are examples of large context effects concerning fairly specific issues - Schuman and Presser, 1981), memory-sampling, context effects are more likely among extremely wide-ranging questions and these probably are among the more common types of context effects.

However, not all wide-ranging questions are readily susceptible to context effects. If people have a predetermined answer to the question that can be directly retrieved rather than having to undertake a general search, then biased memory sampling and therefore context effects are not likely to occur. For example, the inquiry "What is your favorite food?" in theory might require a search of a lifetime of eating experience. However, many people have consciously decided on their "favorite" food and have to report only on that judgment rather than retrieve and evaluate all their memories regarding foods.

Second, ambiguous terms may make questions more susceptible to context effects. In the case of the question on integrated churches

it is likely that the context redefined respondent's understanding of the term "church" and a similar impact may have occurred on the church membership question. Or in the example immediately above people may interpret "food" as referring to main courses, desserts, or five course dinners depending on how prior context makes them understand the vague term "food." These type of context effects depend on the redefinition of question meaning and intent rather than in selective memory sampling as above (Smith, 1989).

Third, demographics seem to be relatively immune to context effects. We did not find a single example of a context effect for a demographic and the literature reports few such effects. However, 2 of our probable 12 effects are factual, behavioral reports and these are as likely to occur as attitudinal effects. (There are more attitudinal effects because there are more attitudinal questions on the GSS.) Demographics differ from the other items in several regards. First, they usually deal with concrete, basic facts that are well-known by respondents. Second, they deal with matters that are well understood by respondents. For most demographics people do not have to do extensive memory searches, but merely go to a discrete point where the information is stored and retrieve it. There is also probably less misunderstanding between respondents and investigators on what is being asked for, since the information - age, marital status, etc. is commonly used and frequently asked for. To some extent standard demographics tend to avoid the memory sampling and definition problems cited above.

This is not to say that 1) there are no problems and no misunderstandings in demographics or 2) context cannot effect demographics, but that demographics tend to be less context sensitive than other questions. (They are relatively immune to context, because respondents understand what is being asked for and that information is readily available.) Also, the above generalizations about demographics apply to simple, basic demographics. Factual states that are difficult to report on or obscure may suffer as readily from context effects as did the items on church membership/integration.

Among the probable context effects on the GSS, we have examples of several of the different types of effects described by various taxonomies and theories of context effects (Tourangeau, et al., 1986; 1988; Smith, 1982; 1989). These include effects acting at the interpretation, retrieval, and judgment stages. We have also identified three factors that contribute to context effects. Context effects are more likely to occur in questions that 1) require wide-ranging, memory searches because the subject covers many relevant memories, 2) access memories that have not been previously organized into a summary evaluation that supplies a simple, direct answer to the question being posed, and 3) utilize ambiguous terms and/or have uncertain intent. These are clearly not the only characteristic of questions that are relevant to context effects (for an extensive list see Tourangeau and Rasinski, 1988), but may be among the most common.

Table 6.1

Statistically Significant Differences Across Ballots by Contextual Groups

(1988 and 1989)

Number Statistically Significant

	SRS		Adjusted (Chi Square * .667)	
1988				
Contextual Groups				
HEF/Sample Frame (9)	2		0	
Start (36)	1	8.3%	0	0.7%
Religion Module (92)	9		1	
Sex Behavior Supplement (7)	0		0	
Other (358)	33	9.2%	13	3.6%
Total (502)	45	9.0%	14	2.8%
1989				
Contextual Groups				
HEF/Sample Frame (9)	0		0	
Start (36)	0	1.6%	0	1.6%
Forbid/Allow Experiments (6)	1		1	
Sex Behavior Supplement (10)	0		0	
Other (381)	35	9.2%	13	2.4%
Total (442)	36	8.1%	14	3.2%

Table 6.2

Difference Across Ballots for Statistically Significant Items in 1988 and 1989

A. Pre-1988 Variables

	1988				1989				Pre-1988			
	A	B	C	Prob.	A	B	C	Prob.	"A"	"B"	"C"	Prob.
MEMCHURH (Yes)	-	.373	.318	.066	-	.360	.295	.029	-	.388	.321	.000
CAPPUN (Favor)	.789	.751	.745	.240	.773	.797	.779	.659	.789	.760	.744	.000*
COURTS												

(Not Tougher)	.105	.088	.117	.444	.040	.025	.020	.114	.089	.075	.121	.000*
5 Satisfactions (LT 10)	-	.238	.321	.032	-	.259	.332	.127	-	.252	.333	.000*
13 Confidences (7+)	-	.108	.173	.035	-	.124	.174	.186	-	.120	.107	.000*
DRUNK (Yes)	-	.343	.410	.070	-	.352	.368	.676	-	.358	.418	.000*

B. 1988 Variables

HELPFUL (Yes)	-	.556	.443	.001*	-	.529	.479	.239	-	.569	.498	.000*
TRUST (Yes)	-	.346	.433	.004*	-	.342	.474	.000*	-	.374	.459	.000*
SATCITY (Great Deal+)	-	.431	.542	.004*	-	.417	.479	.003*	-	.449	.512	.000*
AGED (Good)	-	.393	.499	.003*	-	.393	.459	.033	-	.418	.502	.000*
ANOMIA6 (Agree)	.345	-	.435	.005*	.355	-	.410	.070	.343	-	.403	.000*
SUICIDE1 (App.)	.475	.563	-	.007*	.494	.486	-	.822	.465	.515	-	.000*
RACCHURH (Integrated)	-	.422	.546	.000*	-	.351	.454	.003*	-	.361	.439	.000*
JOBMEANS (# 1)	.471	-	.514	.003*	.529	-	.538	.864	.462	-	.507	.007*
VETYEARS (1+)	.851	.812	-	.003*	.177	.160	-	.720	.818	.815	-	.526
RACOPEN (Favor)	-	.379	.418	.007*	-	.359	.391	.458	.487	.452	-	.026
BUSING (Favor)	.382	.300	-	.008*	.286	.303	-	.551	.213	.273	-	.000*
MEMYOUTH (Yes)	-	.084	.134	.013*	-	.111	.082	.120	-	.106	.091	.057
MEMNAT (Yes)	-	.010	.035	.008*	-	.031	.039	.500	-	.041	.028	.007*
SAVESOUL (Yes)	.443	.427	.524	.005*	-	-	-	-	-	-	-	-

C. 1989 Variables

ABRAPE (Yes)	.828	-	.794	.194	.802	-	.861	.013*	.840	.828	.806	.002*
ALLOWRAC (For)	-	-	-	-	.651	.526	-	.006*	-	-	-	-
ANOMIA7 (Agr.)	.650	-	.710	.049	.620	-	.703	.006*	.667	-	.697	.014*
CONCLERG (Great deal)	-	.172	.241	.025	-	.195	.251	.007*	.331	.278	.307	.000*

CONFINAN												
(Great deal)	-	.253	.296	.302	-	.151	.233	.005*	.274	.228	.298	.000*
HELPFUL (Yes)&	-	.556	.443	.001*	-	.529	.479	.239	-	.569	.498	.000*
LIBATH (Allow)	.354	-	.342	.737	.351	-	.263	.003*	.373	-	.336	.001*
LIBMIL (Allow)	.384	-	.435	.111	.427	-	.340	.003*	.419	-	.395	.073
RACCHURH												
(Integrated)	-	.422	.546	.000*	-	.351	.454	.003*	-	.361	.439	.000*
RELIG16												
(Prt.)	.643	.635	.651	.196	.625	.659	.608	.003*	.649	.637	.665	.084
SATCITY												
(Grt deal +)&	-	.431	.542	.000*	-	.417	.478	.011*	-	.449	.512	.000*
SATFRND												
(V. grt deal)	-	.250	.325	.054	-	.284	.382	.004*	.278	.254	.346	.000*
TRUST (Yes)&	-	.346	.433	.000*	-	.342	.474	.000*	-	.374	.459	.000*
WORDD (Crct.)	.062	.064	-	.328	.026	.072	-	.001*	-	-	-	
WRKSLF												
(Self)	.109	.122	.111	.473	.152	.087	.122	.011*	.126	.124	.123	.959

Notes: Probabilities are based on SRS. If still statistically significant when adjusted for design effects (chi sq. x .667), an asterisk (*) appears. For pre-1988 "A" is 1982 and 1985, "B" is 1983 and 1986, and "C" is 1984 and 1987. &=Also appears on 1988 list.
The mnemonics appearing in capitals are defined in Davis and Smith, 1990. In parentheses are the categories for which the proportions are reported in columns A, B, and C.

Table 6.3

Statistically Significant Differences

	Pre-1988	Dir.	1988	Dir.	1989	88-89 Pooled	Judgement
A. Pre-1988							
CAPPUN	.0000*	O	.240	O	.659	.638	No
COURTS	.0000*	O	.444	O	.114	.185	No
DRUNK	.0000*	O	.070	O	.944	.131	No
MEMCHURH	.0000*	S	.066	S	.029	.005*	Yes
Five Sats.	.0000*	S	.032	S	.127	.0009	Yes?
13 Confid.	.0000*	S	.035	S	.186	.013	Yes
B. 1988 Likely							
AGED	.0000*	S	.003*	S	.033	.0001*	Yes

ANOMIA6	.0000*	S	.005*	S	.070	.001*	Yes
HELPFUL	.0000*	S	.001*	S	.239	.001*	Yes
JOBMEANS	.007*	S	.003*	S	.864	.046	Yes?
RACCHURH#	.0000*	S	.000*	S	.003*	.0000*	Yes
SATCITY#	.0000*	S	.000*	S	.003*	.0001*	Yes
SUICIDE1	.0000*	O	.007*	O	.822	.086	No
TRUST#	.0000*	S	.004*	S	.000*	.000*	Yes

C. 1988 Unlikely

BUSING	.0000*	S	.008*	O	.551	.159	No
MEMNAT	.007*	O	.008*	S	.500	.038	?
MEMYOUTH	.057	S	.013*	O	.120	.526	No
SAVESOUL	---	-	.005*	-	---	---	No
RACOPEN	.026	O	.007*	S	.458	.074	No?
VETYEARS	.526	O	.003*	O	.720	.056	No

d. 1989

ABRAPE	.002*	S	.194	O	.013*	.439	No
ALLOWRAC	---	-	---	-	.006*	---	No
ANOMIA7	.014*	S	.049	S	.006*	.001*	Yes
CONCLERG	.0003*	S	.025	S	.007*	.0004*	Yes
CONFINAN	.0000*	S	.302	S	.005*	.0048*	Yes
LIBATH	.001	S	.737	S	.003*	.021	?
LIBMIL	.073	S	.111	O	.005*	.402	No?
RELIG16	.084	O	.196	O	.003*	.0005*	No
SATFRND	.0000*	S	.054	S	.004*	.0001*	Yes
WORDD	.902	O	.328	S	.0009*	.023	No?
WRKSLF	.959	O	.473	O	.011*	.381	No

#=Also on 1989 list.
*=significant at .05 level after adjustment for design effects
Dir.=Direction: S=Same as 1989; O=Opposite 1989

REFERENCES

Astin, A.W., Green, K.C., Korn, W.S., Schalit, M., and Berz, E.R., The American Freshman: National Norms for Fall 1988, Los Angeles: The Higher Education Research Institute, 1988.

Bishop, G.F., Oldendick, R.W, and Tuchfarber, A.J., "What Must My Interest in Politics Be If I Just Told You `I Don't Know'," paper presented to the American Association for Public Opinion Research, Hunt Valley, Maryland 1982.

Bradburn, N.M., "Response Effects," in Peter H. Rossi and James D. Wright (eds.), The Handbook of Survey Research, 1985, New York: Academic Press.

Cowan, C.D., Murphy, L.R., and Weiner, J., "Effects of Supplemental Questions on Victimization Estimates from the National Crime Survey," Proceedings of the American Statistical Association Section on Survey Research Methods, 1978.

Davis, J.A. and Smith, T.W., General Social Surveys, 1972-1990: Cumulative Codebook. Chicago: NORC, 1990.

Gibson, C.O., Shapiro, G.M., Murphy, L.R., and Stanko, G.J., "Interaction of Survey Questions as it Relates to Interview-Respondents Bias," Proceedings of the American Statistical Association Section on Survey Research Methods, 1978.

Groves, R.M., Survey Error and Survey Costs. New York: John Wiley and Sons, 1989.

Johnston, L.D. and Bachman, J.G., Monitoring the Future: Questionnaire Responses for the Nation's High School Seniors-1975, Ann Arbor: Institute for Social Research, 1980.

McClendon, M.J. and O'Brien, D.J., "Question Order Effects and the Sense of Well-Being," paper presented to the American Association for Public Opinion Research, Buck Hill Falls, Pennsylvania 1983.

Roper, B.W., "The Subtle Effects of Context," The Public Perspective, Vol. 1, 1990, p. 25.

Schuman, H. and Presser, S., Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording, and Order. New York: Academic Press, 1981.

Schuman, H. Kalton, G., and Ludwig, J., "Context and Continuity in Survey Questionnaires," Public Opinion Quarterly, Vol. 47, 1983, pp. 112-115.

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

Smith, T.W., "An Experimental Comparison of Clustered and Scattered Scale Items," Social Psychology Quarterly, Vol. 46, 1983, pp. 163-168.

Smith, T.W., "Can We Have Confidence in Confidence? Revisited," in Denis F. Johnston (ed.), Measurement of Subjective Phenomena, 1981, Washington, D.C.: GPO.

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

Smith, T.W., "Conditional Order Effects," GSS Methodological Report No. 20, Chicago: NORC, 1982.

Smith, T.W., "Counting Flocks and Lost Sheep: Trends in Religious Preference Since World War II," GSS Social Change Report No. 26, Chicago: NORC, 1988.

Smith, T.W., "House Effects: A Comparison of the 1980 General Social Survey and the 1980 American National Election Study,"

Public Opinion Quarterly, Vol. 46, 1982, pp. 54-68.

Smith, T.W., "In Search of House Effects: A Comparison of Responses to Various Questions by Different Survey Organizations," Public Opinion Quarterly, Vol. 42, 1978, pp. 443-463.

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

Smith, T.W., "Some Thoughts on the Nature of Context Effects," paper presented at the First Nagshead Conference Cognition and Survey Methodology, Nags Head, North Carolina 1989.

Smith, T.W., "Timely Artifacts: A Review of Measurement Variation in the 1972-1988 GSS," GSS Methodological Report No. 56, Chicago: NORC, 1988.

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

Tourangeau, R., "Attitude Measurement: A Cognitive Perspective," Hans-J. Hippler, Norbert Schwartz, and Seymour Sudman (eds.), Social Information Processing and Survey Methodology, 1987, New York: Springer-Verlag.

Tourangeau, R. and Rasinski, K.A., "Context Effects in Attitude Surveys." Unpublished NORC Report, 1986.

Tourangeau, R. and Rasinski, K.A., "Cognitive Processes Underlying Context Effects in Attitude Measurement," Psychological Bulletin, Vol. 103, 1988, pp. 299-314.

Tourangeau, R., Rasinski, K.A., Bradburn, N, and D'Andrade, R., "It's Not Just What You Ask but When You Ask It," Unpublished NORC Report, 1988.

Turner, C.F., and Martin, E. (eds.), Surveying Subjective Phenomena, New York: Russell Sage, 1984.