Rotation Designs of the GSS

Tom W. Smith

National Opinion Research Center

University of Chicago

GSS Methodological Report No. 52

February, 1988

This research was done for the General Social Survey project directed by James A. Davis and Tom W. Smith. The project is funded by the National Science Foundation, Grant No. SES-8747227.

Since its inception the GSS employed a rotation design under which most of its items appeared on two out of every three surveys. There are three rotation patterns, so overall the data appeared as in Table 1.

Table 1

Item Appearance on the GSS

1972-1987

Surveys

	1	2	3	4	5	6
Permanent items	X	X	X	X	X	X
Rotation 1	X	X		X	X	
Rotation 2		X	X		X	X
Rotation 3	X		X	X		X

This rotation scheme (designed by Otis Dudley Duncan) allowed the GSS to include more regular items (since each rotating items appeared only 2/3 of the time), but still provided for the regular and reasonably dense repetition of questions. The importance of the rotation scheme increased in recent years, as more items were shifted from permanent to rotating status in order to open up sufficient room for the topical modules.

While this design proved to be a very useful device for both monitoring change and augmenting the content of the GSS, it had the disadvantage of irregularly spacing the data and allowing gaps in the time series. This situation was particularly accute during the 1978-1982 period when NSF did not fund surveys in 1979 and 1981. At that juncture four-year gaps regularly appeared in the data and six-year lapses existed for bivariate correlations between items from different rotations. Even with annual surveys two-year gaps and three-year intervals for bivariate correlations occur.

To eliminate this imbalance in the time series and reduce the length of intervals, we switched in 1988 from an across-survey design as previously used to a split-ballot design.[1] Under this design rotation 1, 2, and 3 would occur across random sub-samples within each survey rather than across surveys (and years). Each sub-sample (known as "ballots") consists of 1/3 of the sample.

Table 2

Item Appearances on the GSS, 1988+ Surveys

	Surveys								
	1			2			3		
	Ballots			Ballots			Ballots		
	A	B	C	A	B	C	A	B	C
Permanent Items	X	X	X	X	X	X	X	X	X
Rotation 1	X	X		X	X		X	X	
Rotation 2		X	X		X	X		X	X
Rotation 3	X		X	X		X	X		X

Table 2 shows how ballots now take the place of surveys (years). Permanent items appear on all ballots of all surveys. Items in rotation 1 appear on ballots A and B on each survey; items on rotation 2 appear on ballots B and C; and items on rotation 3 on ballots A and C. In effect ballot A contains those items that would have appeared under survey (year) 1 of the

[1]The revised rotation plan evolved from suggestions made by James A. Davis to the GSS Board of Overseers in May, 1984 (Davis, 1984). In a series of consultations over the next year between Davis and Smith, the Overseers, and several colleagues (Howard Schuman, Roger Tourangeau, William Kruskal, Otis Dudley Duncan, and Seymour Sudman), the present rotation design was developed (Alwin, 1986). This plan was submitted to NSF as part of the five-year renewal proposal of the GSS (Davis and Smith, 1985).

old rotation across survey scheme (Table 1), ballot B represents survey (year) 2, and ballot C survey (year) 3. As one can see by comparing surveys (years) 1-3 in Table 2, the content of the core GSS no longer varies across surveys (years), but remains fixed.

In terms of appearances permanent items are not affected by this switch. They continue to appear on all cases for all surveys. Rotating items now appears on all surveys and are asked on each survey of 2/3 of respondents. Over a three-year cycle the same number of respondents are asked the "rotating" items as before (3,000), but instead of coming in two segments of 1,500 each from two surveys, they appear in three segments of 1,000 each from three surveys.

Table 3

Number of Items on the 1988 GSS by Rotation Status

	Ballots		
	A	B	C
Replicating Core			
Permanent	169	169	169
Rotation 1	92	92	0
Rotation 2	0	94	94
Rotation 3	55	0	55
	316	355	318
Supplemental Items			
Topical Module (Religion)			
Upgrades	0	0	27*
Additions	70	70	70
ISSP	58	58	58
AIDS Questions	27	27	27
	471	510	500

*Religious items normally appearing on rotation 1 that will also be asked on ballot C.

Table 3 shows how the items in 1988 fit into the split-ballot design. There are 169 permanent items that appear on all forms, 92 rotating items that appear on ballots A and B, 94 rotating items that appear on ballots B and C, and 55 rotating items that appear on ballots A and C. Together these make up the replicating core of the GSS and they will appear in the same fashion over the next four surveys. Altogether there are 316 core items on ballot A, 355 on ballot B, and 318 on ballot C. The bottom half of Table 3 shows the appearance of supplemental items. These items are not part of the replicating core and are not governed by the rotation design or ballots. Normally these items appear in a single year only, although supplemental items may repeat in different surveys.[2] One special feature of the 1988 design is that the topical module on religion not only added 70 new items, but also upgraded 27 religious items that are part of the replicating core. These items are part of rotation 1, regularly appearing on ballots A and B. In 1988 they were added on ballot C so that they would be were asked of all respondents just like the new religion items in the topical module

There are several major advantages of the split-ballot design. First, we will have a better sampling of time. Since all items will be measured each year, there will no missing observations and all intervals will be one year (for both permanent and "rotating" items). This will, in particular, improve the tracking of change triggered by episodic events (e.g. the Soviet invasion of Afghanistan) and the identification of "turning points", such as the reversal in the pro-legalization trend on marijuana in the late 1970s.

---

[2]For example, a substantial part of the 1985 network module was repeated as part of the socio-political participation module in 1987 and part of the 1985 ISSP role of government module will be repeated in 1990.

Second, it will be easier to judge the comparative rate of change across items from different rotations. As Table 1 indicates an item on rotation 1 starts in year 1 and ends in year 5 while a item on rotation 2 starts in year 2 and ends in year 6. One could of course compare years 2 through 5 on both rotations (at the cost of "losing" the readings in years 1 and 6), but if one also wanted to compare items from rotations 3, the problem becomes unsolvable since items from all three rotations neither start nor end in the same year. On the split-ballot design all items of course appear in all years.

Third, it is simpler to apply econometric, time-series analysis techniques to evenly spaced items without gaps in the time series. In particular the split ballot technique facilitates the linkage of other annual data series to the GSS trends, allowing the investigation of the relationship between public attitudes and aggregate time series on such matters as unemployment and inflation, criminal victimization, and government expenditures.

Fourth, since we designed the three ballots to duplicate as closely as possible the across-survey rotation cycle, we can now test for context effects to see if there have been any contextual distortions in the time series. (On efforts to avoid this problem under the old design see Smith, 1986). Context effects are discussed in more detail below.

There are, however, also several disadvantages associated with the split-ballot design. Probably the chief disadvantage of the split-ballot design is that each single survey reading has more sampling variability than under across-survey rotation (two sigma limits for a proportion of .5 rise from .032 to .039, assuming a design effect of 1.5). Of course, for the missed year under the old design there is no information (and an infinite

sampling variation), while the split-ballot design has a two sigma level of .039 every year.

Furthermore there are two analytic approaches that minimize or even eliminate the loss in precision. A rolling average type of approach (i.e. an averaging of years 72-74, 73-75, 74-76, etc.) would have similar precision under both across-survey and split-ballot designs. Perhaps even more promising would be a combination of trend analysis and pooling. If we pooled over three year intervals and then compared the successive pooled readings (e.g. 1972-74, 1975-1977, 1978-1982, 1983-1985, etc.), we would have approximately 3,000 cases for each of the pooled time points under either method. This type of combination of pooled time series analysis would be especially useful for examining changes among subgroups, for averaging over random sampling variability from survey-to-survey, and for studying longer range change rather than annual trends.

Second, the switch in designs complicates the equal representation of time in cumulative GSS analysis. The old rotation plan also created some problems along these lines and the impact of the across-survey and split-ballot designs as well as other factors are discussed in Appendix 1.

Third, the split-ballot design is more expensive than the across-survey approach. Three separate questionnaires have to be prepared, three sets of show cards, three data entry programs, three sets of cleaning specifications, and so forth. In addition interviewer training must be increased since they will have to master three different (although similar) questionnaires. At present we estimate an increased cost of 7%. However, we anticipate that in future years part of the higher costs will be recouped. Since the replicating core will not change from year to year, we have designed the three ballots so that the first two-thirds of each ballot (essentially the

part prior to the topical module on religion) will remain completely unchanged. Thus there should be no revisions in the questionnaires, coding instructions, data entry programs, or cleaning specifications for that part of the survey. This savings can be achieved, however, only if we rigorously resist making changes in this part of the instrument. This means that certain constraints are placed on the design and implementation of methodological experiments (see Appendix 2).

Fourth, a potential disadvantage has to do with the possible disruption in our time series by the introduction of new context effects. If we had recreated de novo the three rotational groupings, we would have changed both the possible associational relationships that could be investigated (mainly what three way and higher order relationship that could be tested) and the order in which the questions appeared. To avoid such context changes and other alterations, we instead maintained the three rotations as previous assigned. Because of this, the three ballots duplicate as closely as possible the last three years of the GSS: A is 1985, B is 1986, and C is 1987. Or, to put it another way, they represent what the next three years would have looked like if the split-ballot design had not been adopted.[3] After the 1988 data are collected we will carry out an extensive analysis of differences across ballots to see if any of the across time variation observed in the past might have been due to the across-survey rotation design rather than to true change. While the conversion of the old rotations directly into the new split-ballot design avoids the problem of new context effects beings created and allows for the investigation of context effects under the old across-

[3]"Duplicate as closely as possible" means that all of the same questions appear in the same order except for the deletion of some old experiments, a couple minor changes to regularize rotation patterns (e.g. with WORDSUM), and changes where the topical module is inserted.

survey design, it in turn has some costs. In general one wants to keep related items on the same rotation. Three rotating items on the same rotation produce a 1,000 case three-way cross-tabulation on each survey or, to think of it in a slightly different way, one could construct a 1,000 case three-variable scale. If one of the items appeared on a different rotation, the cross-tabulation or scale of the three falls to 500 cases per survey and if the three items appeared on three different rotations, then no three-variable crosstabuation or scale would be possible. While this principle was generally followed under the across-survey rotation design, a second competing desire to have some items on a topic appear every survey (year) meant that some items were intentionally spread across rotations. Consider the five racial attitudes in Table 4. They appear on all three rotations. Since under the split-ballot design these items now all appear each year, the original rationale for their spread across rotations has disappeared and the principle of asking related items on the same rotation argues for their consolidation.

Table 4

Rotation of Selected Racial Items

Items	Years		
	1982	1983	1984
Rotation 1 RACMOST, RACPRES	X	X	
Rotation 2 RACOPEN		X	X
Rotation 3 RACMAR, RACPUSH	X		X

To do so would however disrupt the established order of items and change the combination of variables that could be compared in three way and higher analyses. To avoid this problem we opted to maintain the existing rotations despite non-optimal grouping of some variables.[4]

Finally, users will have to pay more attention to what variables can be utilized in multivariate models than previously. Under the across-survey design, it was more obvious what variables appeared together. One looked up the years that a variable appeared in and if they were asked in the same year, one could carry out the analysis without paying any explicit attention to what rotation it was on. Under the split-ballot design one must determine what rotation variables are on (or on what ballots they appear) to know whether 3+ variable models are possible. No changes occur in the possible models under the split-ballot design, but one will explicity have to consider the "rotation" design to know what is possible and what is not.

The rotation design of the GSS permits the inclusion of more variables at the cost of a restriction on the inter-item analytical possibilities and some added complexity. Overall, we have found this trade-off to be beneficial. The new split-ballot design does not fundamentally alter the situation under the across-survey design, but does change how time is sampled. We believe that the gains in having annual, uninterrupted readings of variables unweighs the increase in sampling variability (for years that

---

[4]The old rotation scheme, this new scheme, and any similar use of sub-samples naturally limits the possible analysis of inter-item associations. Only a design under which all items were asked of all respondents would permit the full, unrestricted analysis of relationships. Such a design in turn minimizes the number of variables that can be included and thus analyzed. Conversely, a design with multiple sub-samples which each asked a unique set of questions would minimize the study of inter-item relationships (since no cross-tabulations across sub-samples would be possible), which maximizes the number of variables covered.

would have been covered under the across-survey design), the higher costs, and other disadvantages. We will be monitoring the situation over the next four years to see if this evaluation proves to be correct.

Appendix 1: Sampling Time - The Cumulative File

The GSS rotation design as well as other factors affect how time is represented when the GSS is used as a pooled or cumulative file combining all years together. In addition to being a spatial sample of the United States, the GSS is also a temporal sample of the years 1972 to the present. This sample of the United States over the period 1972-1987 can be thought of as stratified by time (years), much as it is stratified by space (region and rural/urban). If we take years as equal units of time, then each year should sample an equal number of people. (Since the population is actually growing, the sample fraction is slightly increased each year. One might argue that a constant proportion of the population should be sampled each year which would lead to a slow and steady increase in the sample size reflecting the growing population. We however prefer to think of each year as a equivalent unit that should have equal representation.)

Several factors have created deviations from this goal. First, while the GSS aims for 1,500 completed cases each year the actual number has averaged 1,512 (from 1972 to 1987, excluding oversamples) and has ranged from a low of 1,466 in 1987 to a high of 1,613 in 1972. While these deviations could be adjusted for, we feel that they are small enough to be ignored for most analytic purposes.

Second, because no surveys were funded in 1979 and 1981 these years have zero representation. This not only means that these two years are excluded from any analysis, but that any consecutive, temporal period covering those years is underrepresented. Thus, a pooled analysis for 1972-1986 has five surveys and 7,590 cases for 1972-1976, three surveys and 4,530 cases for 1977-1981, and five surveys and 7,582 cases for 1982-1986. A permanent item (e.g. a questions that is asked of all respondents on all GSSs) is underrepresented from the middle period by 31.1%.

Third, not all items have been asked from the beginning to the end of the series. Some items have been dropped and a larger number have been added. This presents little problem when using the pooled file since the analysis is by necessity simply restricted to the years that have been covered. In conjunction with the rotation pattern across years and the switch in rotation designs, the late start or early end of time series can however slightly complicate the equal representation of time. Adding a rotation item after the start of its rotation (i.e. in the middle rather than at the beginning of its three-year cycle) or dropping an item before the finish of the rotation, means that the design for representing time was not fully implemented. Usually any problem from this situation might be avoided by simply referring to the actual years covered. In some circumstances and from some perspectives this may contribute to imbalances in the representation of time. Consider an item that was to follow the rotation

1 2 3 4 5 6 7 8 9
X X 0 X X 0 X X 0

(i.e. asked in years 1 and 2, skipped to year 3 and so forth). If the items was actually added in year 2 and then dropped after year 7 the pattern would

2 3 4 5 6 7
be X 0 X X 0 X. By two year periods 2-3 and 6-7 would have only one reading each and thus would be underrepresented compared to years 4-5 with two readings. Other complications in comparing item from different rotations can also be increased by these incomplete rotation cycles.

Fourth, numerous GSS items have been subjected to methodological split-form experiments. Under the typical experiment an item appears in its standard form (or order) on only one-half of the sample (or occasionally or two-thirds) while a variant wording (or order) occurs on the other half (or third). That means that the standard item is underrepresented for the year in question. Sometimes, when the experimental comparison reveals no statistically significant difference between the two forms, one might be able

to preserve the balance by combining the standard and variant versions. (This procedure can be justified, but is subject to potential problems. For example, two variables may not show any marginal differences, but may correlate with a third variable in a different way or at least to a significantly different degree.)

Fifth, many items follow a rotation plan. From 1972 and 1987 that meant that the item appeared on two out of three years. In effect, the design was to systematically sample time by asking the rotating items in either years 1 and 2, 2 and 3, or 1 and 3. Thus across three-year cycles time is represented equally. In 1988 a new rotation scheme was adopted under which rotating items would appear on two-thirds of the cases each year. Over a three-year cycle the item will be asked of the same number of respondents under both designs. Thus the comparison of three years blocks will represent those years equally although the distribution of respondents within the three-year cycle will differ. However, when comparing less than complete cycles the new rotation design will underrepresent the split-ballot years versus the across-survey years. For example, a comparison of 1986-1989 will draw 2,936 cases for an item asked in 1986 and 1987 (i.e. an item on rotation number 1), while for 1988 and 1989 the projected number of cases is 2,000.

Sixth, particularly in recent years some items have been switched from permanent items to rotating items. Either under the old, across-survey approach or under the new, split-ballot approach this means that the temporal balance is upset. For example, the five-item satisfaction battery was asked annually from 1982 to 1984 since it was a permanent item. In 1985 it was switched to being a rotating item. It was off-rotation in 1985 and asked in 1986 and 1987. As a result of this shift it was asked of 4,578 respondents in 1982-84, but only 2,936 respondents in 1985-87, thereby overrepresenting the prior period.

Finally, several of the above factors can work together to alter the balance across time. For example, in 1984 the confidence items were part of an experiment and were asked of two-thirds of respondents. In 1985 they switched from permanent to rotating status and did not appear that year. Thus confidence in underrepresented in 1984 and, compared to earlier three-year cycles (e.g. 1973-1975) underrepresented in 1985-87.

In brief, we see that in many particular instances items are not equally distributed over time so that the sampling of time can be considered biased. Does this make a difference? Potentially yes. If the distribution of an item or its association with other variables is different during certain years and if these years are under or overrepresented for the reasons cited above, then the pooled data analysis would be biased. Let's take a simulated example. The proportion saying that spending for welfare is too little was 21% in 1973, 23% in 1974, 25% in 1975, 14% in 1976, 13% in 1977, and 13.5% in 1978; a rather pronounced step function. Over this period 18% supported more spending for welfare. Now if a switch to rotation had been made after the first three years and if the middle year (1977) in the second three-year cycle had been skipped, then the overall average would have been 19%. This one percent increase can be considered as a bias resulting from the unequal sampling of time (specifically the underrepresentation of the 1976-78 period compared to 1973-75). The magnitude of the bias in rather modest however and an inspection of many series indicates that the average disortion is less than the one percent in this example.[1] While the analyst should carefully consider

---

[1]The bias will be greatest when the imbalance in sample size across years is greatest and the difference in the observed values of a variable is largest. This imbalance seems to greatest if two years are being combined in which one year asked the question of a full sample and the other year asked the question on an experimental half sample. The largest bias that I could find was comparing the level of trustworthiness (TRUST) in 1983 and 1984. In

whether bias from the unequal representation of time exists and whether any weighting should be applied to adjust for this bias, in general the distortion will probably not necessitate any correction.

1983 an experiment was carried out and TRUST was asked of a half sample (N=802) while in 1985 a full sample (N=1465) answered the question. The proportion saying people could be trusted varied significantly across these years (.369 in 1983 vs. .478 in 1984). Using the raw Ns the pooled proportion was .439, but giving equal weight to the years gave a pooled proportion of .424, .015 lower. Of course if a large interaction with time occurs, one might better study the change rather than pool the data. This latter approach would still be legitimate however. It would not be much different than "pooling" results from different regions in a cross-sectional survey when the regional differences were significantly different.

Appendix 2: Methodological Experiments Under the Split-Ballot Design

In addition to the three ballots that are used for the new split-ballot design, the GSS will continue to use experimental forms that will be drawn independently of the ballots. In 1988 there will be two forms (X and Y) and these random half-samples will carry out the same experiments in each ballot. The new split-ballot design does however have an impact on the design and implementation of experiments. In order to achieve the cost sharing alluded to in this report, we do not want to disrupt the order of questions with the insertion of experiments within the fixed, front pages of each ballot. How experiments can be handled depends on what type of items are involved and what type of experiment is being carried out. First, new items being added only for experimental purposes would appear near the end of whatever ballot they were scheduled to appear on, after the fixed, front part of the questionnaires. Second, experiments involving new items being added as part of the topical or international modules would appear as part of these respective sections (the topical module appearing near the end of each ballot and the international module being a self-completed form administered after the main questionnaire).

Third, for a "rotating" item from the replicating core, the standard version would appear on the two ballots in the same position as regularly scheduled. On the third, off-ballot the standard and variant conditions would appear on forms X and Y near the end of that ballot. Analysis could either compare the pure experiment (forms X and Y on the off-ballot) or use a less controlled comparison of the standard version (on the two scheduled ballots and the X form of the off-ballot with the variant version on the Y form of the

off-ballot).[1] The limitation of this approach is that it yields a "pure" experiment with only 250 cases for each version and a less rigorous experiment with 1,250 cases for the standard version and 250 for the variant version. These designs would permit the identification of experimental variations that were moderate-to-large in size. To test for smaller effects, one would want to repeat the experiment across two or more surveys and then either compare the results (Schuman and Presser, 1981) or pool the experiments treating them as one large experiment of 1,000 cases over two years, or 1,500 cases over three administrations.

Fourth, for permanent items the situation is the most difficult. Without disrupting the content of the fixed portion of the questionnaire the one strategy would be to repeat the variant and standard version on the X and Y forms of one or all of the ballots. In addition to getting the simple experimental comparison across forms on the standard and a variant version, one would also pick up a test/retest type measure of reliability on the X forms. This reliability measure on the standard version could also be compared to the standard-variant "reliability" that could be calculated on the Y form. We actually anticipate little need to utilize such a design since almost all of the permanent items are demographics and none of these have been experimentally examined over the past 15 surveys. Still it is a possible option.

There are however a few types of experiments that are either seriously ruled out or compromised by the split-ballot constraints. For example, if one wanted to test rapport vs. fatigue effects on sensitive questions by asking questions at the start of the interview vs. at or near the end, this would be

---

[1] If the standard version of the variable did not vary across the three ballots, this approach might be used.

probably be impossible to carry out on the GSS. Similarly, an experiment that necessitated that no prior questions had been asked about some permanent item (e.g. no prior occupational questions) could not readily be adopted.

In brief, the use of split-ballots to replace the across-survey design creates some constraints in the design and execution of methodological experiments. In general, however, experiments can be designed that conform to these constraints and still allow us to test the effects we wish to explore.

References

Alwin, Duane, "Report of the Board of Overseers of the General Social Surveys, 1983 to 1986," Submitted to the National Science Foundation and the National Opinion Research Center, November, 1986.

Davis, James A., "Design and Redesign of the GSS," Memo to the GSS Board of Overseers, May 7, 1984.

Davis, James A. and Smith, Tom W., "A Proposal to Continue 'A National Data Program for the Social Sciences' (The NORC General Social Survey)," Submitted to the National Science Foundation, July, 1986.

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

Smith, Tom W., "Unhappiness on the 1985 GSS: Confounding Change and Context," GSS Technical Report No. 56. Chicago: NORC, 1986.