

Wording Effects on the National Spending Priority Items across Time, 1984-2004

Tom W. Smith

NORC/University of Chicago

September, 2006

GSS Methodological Report No. 107

Since 1984 the General Social Survey (GSS) has carried out wording experiments involving its national, spending, priorities items. The battery asks the following:

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

This question has been on all GSSs since 1973. In 1984 there were three versions of the traditional 11-item scale, each administered to a random third of the sample. These were: 1) the standard wordings that had been used since 1973 (e.g. NATEDUC), 2) a variant with all terse wordings (the Y questions, e.g. NATEDUCY), and 3) a variant with all verbose wordings (the Z questions, e.g. NATEDUCZ). For the full-wordings of all items see Appendix 1 or Davis, Smith, and Marsden (2007). The standard wordings blended together terse versions that only mention a spending area (e.g. Foreign aid, Space exploration program) with verbose versions that indicated an area and promised some improvement or accomplishment (e.g. Improving and protecting the environment, Solving the problems of the big cities). The terse-only version generally kept the terse version used in the standard items and replaced verbose wordings with terse wordings (e.g. Improving and protecting the environment being changed to The environment). The verbose-only version generally retained the verbose wordings among the standard wordings and replaced terse wordings with verbose wordings (e.g. Space exploration with Advancing space exploration). One of the verbose-to-terse adaptations also dealt with the problem that the standard crime item (Halting the rising crime rate) was based on the often counter-factual idea that the crime rate was increasing by eliminating that assertion in both the terse wording (Law enforcement) and the verbose variant (Reducing crime). The all-verbose, Z-version was used only in 1984, but all subsequent GSSs (1985-2006) continued the split-ballot administration of the standard and Y-versions, each being fielded on half the sample.

In addition, four new items were added to the scale making a total of 15 items starting in 1984 (Highway and bridges, Social Security, Mass transportation, and Parks and recreation). These items were asked in only terse versions, appeared after the standard, 11-items, and were not experimentally manipulated. Subsequently, other items have been added to the scale: Assistance for children starting in 2000 and Supporting scientific research beginning in 2002. These items were added to the end of the list and were not asked with variant wordings. Thus, the current scale has 17 items, 11 of which appear in two, experimental versions (standard wordings and terse-only wordings) and 6 that follow the initial 11 and appear in only one version.

The initial 1984 experiments revealed a number of moderate-to-large wording differences that have been analyzed in Smith (1987) and Rasinski (1989). Consistent with the general GSS practices of maintaining consistent measurement across time (Smith, 2006) and of replicating experiments over time (Schuman and Scott, 1989), and specifically because several of the spending areas did show notable differences in support

across the wordings (Smith, 1987; Rasinski, 1989), the standard items and the terse-only/Y versions have been asked on all GSS from 1984 to the present.

Table 1 shows the results for the experiments on the 11 items for the 15 GSSs conducted in 1984-2004. Overall, these represent about 29,300 cases (exact Ns varying between 29,306 and 29,374 due to missing values). The percentages in the first columns are the difference in the percent saying Too little is being spent on the standard version minus the percent saying Too little on the Y variant. Don't Knows are retained in the base. The second columns report the probability that the difference is statistically significant. The probability levels are based on the four categories (Too Little, About right, Too much, and Don't Know), not just the differences in the Too little percentages. The bottom two rows have the cumulative percentage difference and probability level and the cumulative sample size.

Overall, there is a great deal of consistency in the direction and magnitude of the wording effects. Eight of the 11 items show no statistically significant interaction of question wording with time.¹ Two of the items (The Space Exploration program/Space exploration and The military, armaments, and defense/National defense) show very small average differences (respectively 0.1 points and -0.7 points) and these are not statistically significant even in the large cumulative sample.

Three of the other items also have small, but statistically significant, average differences (+1.0 points for Improving and protecting the nation's health/Health; -0.7 points for Foreign Aid/Assistance to other countries; and -1.3 points for Improving and protecting the environment/The environment), but no statistically significant interaction with time. Health and Foreign aid also show no discernable pattern, but Environment does show possibly systematic differences. In six of the first seven surveys the verbose version shows more support for environmental spending, although only one of the differences is even of borderline statistical significance. In the last eight surveys the terse version has higher support levels, although this is statistically significant in only one year. The interaction is also surprising since across all comparisons the verbose versions regularly generate more support for spending than the terse versions do (Rasinski, 1989). However, since the interaction is not statistically significant, this is not considered further.

Showing larger and statistically significant differences in the cumulative file, but with no statistically significant variation in year-to-year differences, are Improving the nation's education system/Education (+4.0 points), Improving the condition of blacks/Assistance to blacks (+7.0 points), and Welfare/Assistance to the poor (-43.7 points).

That leaves three items with some possibly meaningful temporal variation in the question-wording effects. First, for crime (Halting the rising crime rate/Law enforcement), the standard, verbose version always receives more support (averaging +10.2 points), but the edge diminishes after 1994 with five of the six smallest differences coming in this later period.

Third, for spending for big cities (Solving the problem of the big cities/Assistance to big cities) the standard version always garners much more support than the variant (averaging +26.3 points), but the difference is slightly greater than usual in 1989-1998 than earlier or later and especially high in 1993-1996 (30.1-32.2 points).

¹ Based on ANOVA analyses using several different recodes of the spending priority items.

Finally, the standard item, Dealing with drug addiction, generally gets more support (averaging +6.4 points) than does the variant, Drug rehabilitation. This edge however disappears in 1990 and 1991 (-0.1 and -0.6 in those years) and is generally smaller after this period than before (+9.9 points before vs. +6.0 points afterwards)

These interactions do not appear to follow any general pattern. First, the average difference across all 11 items does not change much year-to-year and the little fluctuations do not show any discernible trend (see last columns in Table 1). Second, there does not seem to be any clear substantive connection across the areas experiencing and not experiencing question-wording interactions. Third, the results do not appear related to any systematic linguistic differences such as verbose, promising items generally losing their pro-spending edge over time.

That suggests that the interactions in wording effects may be due to specific changes relating to the particular words used and/or alterations in how the areas are perceived by people. Perhaps the two most plausible explanations can be hypothesized for the topically-related areas of crime and drugs. First, the decline in the advantage of the Halting the rising crime rate over Law enforcement may have come because more people realized the factual inaccuracy of the former during the appreciable decline in crime during the 1990s. Certainly support for spending fell appreciably during the 1990s as the crime rate dropped (Smith, 2005). The decline was greater for Halting than for Law enforcement and this could possibly derive from more people realizing the inaccuracy of the “rising” assertion. Alternatively, Law enforcement may have taken on a more positive valence over time.

Second, the drop in the edge of Dealing with drug addition vs. Drug rehabilitation over time (and its temporary disappearance in 1990/91) could be due to more people seeing drug rehabilitation as an effective method for dealing with illegal drug use. Alternatively, perhaps the phrase “deal with” has taken over less positive connotations over time, perhaps associated with the catch phrase “deal with it.” The data do show that the closing gap resulted from a large decline in support for Dealing with drug addition from 1989 to 1991 and only a small decline for Drug rehabilitation. Furthermore, Timberlake, Rasinski, and Lock (2001) have shown that the drug, wording effect is not constant across social groups, being larger for example among conservative groups. Thus, changes in the policy and/or politics of drug control could alter the magnitude of the wording effect.

Unfortunately, there is no easy way to test these ideas about possible changes in language use.

Overall, wording effects appear to be quite stable across 20 years of the 15 replicated experiments. This is consistent with results from the handful of measurement experiments that have been replicated across time (Hippler and Schwarz, 1989; Schuman and Presser, 1984; Schuman and Scott, 1989). For eight of the 11 comparisons there is no statistically significant interaction with time. For the three showing statistically significant interaction, the variations are small. The wording effects for City and Crime are always in the same direction and statistically significant in 29 of 30 comparisons. For City only three of the annual differences were more than +/- 4 points from the cumulative average. For Crime only two years varied by +/- 4 points from the overall difference. Drug is the only topic showing a statistically significant interaction in which signs change (negative in two years), but these reversals are not statistically significant. While not

large, the differences for these variables are statistically significant and do show a clustering across years that suggest systematic rather than random variation. This cautions researchers to consider that even well-established measurement effects may not be constant across time.

References

- Davis, James A.; Smith, Tom W.; and Marsden, Peter V., General Social Survey Cumulative Codebook: 1972-2004. Chicago: NORC, 2005.
- Hippler, Hans-J. and Schwarz, Norbert, "Not Forbidding Isn't Allowing: The Cognitive Basis of the Forbid-Allow Asymmetry," Public Opinion Quarterly, 50 (1989), 87-96.
- Rasinski, Kenneth A., "The Effect of Question Wording on Public Support for Government Spending," Public Opinion Quarterly, 53 (1989), 388-396.
- Schuman, Howard and Presser, Stanley, Questions and Answers in Attitude Surveys. New York: Academic Press, 1981.
- Schuman, Howard and Scott, Jacqueline, "Replication of Two GSS Question Form Effect Experiments: 1974 and 1982," Sociological Methods and Research, 17 (1989), 398-408.
- Smith, Tom W., "Formulating the Laws of Studying Societal Change," Version 2.2. Chicago: NORC, 2006.
- 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 (1987), 75-83.
- Smith, Tom W., "Trends in National Spending Priorities, 1973-2004," NORC report, 2005.
- Timberlake, J. M., Rasinski, K.A., & Lock, E.D. (2001) Effects of conservative sociopolitical attitudes on public support for drug treatment spending. Social Science Quarterly, 82 (1), 184-196

Table 1

Differences on the Question-Wording Experiments, 1984-2004

	SPAC %D	SPAC Prob	ENVIR %D	ENVIR Prob	HEAL %D	HEAL Prob	CITY %D	City Prob
1984	1.9	0.134	2.3	0.051	2.3	0.690	25.7	0
1985	1.4	0.558	-4.0	0.089	2.3	0.090	21.6	0
1986	1.9	0.340	1.9	0.450	-0.3	0.088	28.3	0
1987	0	0.010	2.1	0.630	1.3	0.181	18.3	0
1988	-1.7	0.135	0.6	0.458	-2.2	0.118	24.5	0
1989	0.6	0.649	3.4	0.172	-2.5	0.004	27.2	0
1990	0.4	0.839	0.8	0.326	5.3	0.036	28.0	0
1991	-0.6	0.668	-0.4	0.159	0.9	0.607	26.1	0
1993	1.9	0.030	-3.0	0.606	1.8	0	30.7	0
1994	0.4	0.404	-2.5	0.139	1.5	0.001	32.2	0
1996	-1.5	0.533	-3.2	0.257	3.7	0.027	30.1	0
1998	-1.0	0.571	-1.8	0.139	-0.8	0.011	26.5	0
2000	1.1	0.848	-1.6	0.570	2.3	0.008	21.6	0
2002	-0.5	0.471	-3.9	0.002	0.6	0.005	25.7	0
2004	0.3	0.951	-1.2	0.624	-0.4	0.020	22.5	0
Cum.	0.1	0.336	-1.3	0.0002	1.0	0.0000	26.3	0.0000
N		29374		29345		29334		29306

YEAR	CRIME %D	CRIME Prob	DRUG %D	DRUG Prob	EDUC %D	EDUC Prob	RACE %D	RACE Prob
1984	12.8	0.0004	16.5	0	-0.1	0.030	10.6	0.0001
1985	6.3	0.068	8.4	0.0008	-4.8	0.030	5.1	0.009
1986	12.9	0	5.1	0.038	-5.9	0.028	11	0
1987	16.7	0	8.4	0.001	-5.7	0.175	8.5	0.003
1988	13.9	0	11.5	0	-4.6	0.041	9.3	0
1989	11.9	0	9.4	0.002	-5.1	0.006	5.1	0.014
1990	14.0	0	-0.6	0.349	-1.5	0.912	8.1	0.004
1991	10.2	0.0004	-0.1	0.262	-2.7	0.638	3.6	0.048
1993	12.3	0	5.2	0.152	-3.8	0.395	12.1	0
1994	12.3	0	7.5	0.0004	-1.6	0.270	6.7	0
1996	10.1	0	5.6	0.001	-6.6	0.0007	8.1	0
1998	7.3	0.0003	5.7	0.001	-3.6	0.177	7.7	0
2000	8.7	0	6.6	0.001	-4.3	0.055	3.1	0.001
2002	8.6	0	6.2	0	-3.1	0.261	6.4	0
2004	4.6	0.0001	5.4	0.002	-5.7	0.002	5.6	0
Cum.	10.2	0.0000	6.4	0.0000	-4.0	0.0000	7	0.0000
N		29314		29324		29343		29275

Table 1 (continued)

YEAR	DEF %D	DEF Prob	FORAID %D	FORAID Prob	WFARE %D	WFARE Prob	AVERAGE %D
1984	-0.6	0.773	0.6	0.068	-38.3	0	10.2
1985	-1.4	0.353	-0.4	0.304	-44.5	0	9.1
1986	0.1	0.045	1.7	0.184	-39.0	0	9.8
1987	-4.4	0.284	2.9	0.057	-46.8	0	10.5
1988	-1.9	0.562	0.5	0.137	-45.0	0	10.5
1989	-0.8	0.942	-1.0	0.218	-43.2	0	10.0
1990	0.8	0.589	-0.2	0.308	-44.0	0	9.4
1991	1.2	0.227	0.8	0.773	-42.5	0	8.1
1993	-1.5	0.272	1.6	0.025	-46.3	0	10.9
1994	-0.2	0.523	0.3	0.276	-44.8	0	11.0
1996	0.4	0.988	0.3	0.001	-39.6	0	9.9
1998	0.0	0.951	0.6	0.0002	-44.0	0	9.0
2000	-2.0	0.560	0.3	0.001	-42.0	0	8.5
2002	-3.5	0.012	-2.3	0	-45.0	0	9.6
2004	1.5	0.719	1.4	0.057	-45.2	0	8.5
Cum.	-0.7	0.087	0.3	0.0000	-43.7	0.0000	
N		29330		29315		29326	

Note: 0 in Prob column stands for 0.0000. %D=difference in percentage points between % Too low on standard version minus % Too low on Variant/Y version.

Appendix 1: National Spending Priority Items

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

Items covering 11 areas have been asked in every GSS since 1973 (Space Exploration Program, Improving and Protecting the Environment, Improving and Protecting the Nation's Health, Solving the Problems of the Big Cities, Halting the Rising Crime Rate, Dealing with Drug Addiction, Improving the Nation's Education System, Improving the Condition of Blacks, The Military, Armaments, and Defense, Foreign Aid, Welfare).

Since 1984 experiments have been conducted and 11 alternative wordings for the original spending items have been asked on a random sub-sample (Space Exploration, The Environment, Health, Assistance to Big Cities, Law Enforcement, Drug Rehabilitation, Education, Assistance to Blacks, National Defense, Assistance to Other Countries, Assistance to the Poor).