A COMPENDIUM OF TRENDS ON GENERAL SOCIAL SURVEY QUESTIONS

by

Tom W. Smith

with assistance from

Guy J. Rich

GSS Technical Report No.:15

National Opinion Research Center

August, 1979

This work was carried out for the General Social Survey Pro~ect which is funded by the National Science Foundation, SOC-7703279.

INTRODUCTION

Selection and Coverage

The compendium contains a selected summary of trends on General Social Survey items. For non-demographics the trend series contains all data points known and available to the GSS staff that have an exact or equivalent national usage. [An exact usage is a verbatim duplicate. An equivalent usage has minor differences in wording such as the alternation of a article or pronoun, or some other trivial change. Some close variations are also included in the series. mese variations are noted when they were judged suitable for inclusion. In general, the practi~e was to exclude all but the most insignificant of variations.] It excludes variant wordings and data from local or restricted populations (e.g., the Detroit Area Study, by the University of Michigan or a national sample of Catholics by NORC). A selective list that also cites variant wordings and data points covering restricted populations can be found in Appendix M: Previous Usage in James A Davis, Tom W. Smith, and C. Bruce Stephenson, General Social Surve s 1972-1978: Cumulative Codebook (Chicago: NORC, 1978). Other more general guides to previous usages and other time series include Jessie C. Southwick, ed., Survey Data for Trend Analysis: An Index to Repeated Questions in U.S. National Surveys Held by the Roper Public Opinion Research Center, (Williamstown, Mass.: Roper Public Opinion Research Center, 1975), "A Continuity Guide to the American National Election Surveys of the Center for Political Studies, 1952-1974," (Ann Arbor: Institute for Social Research, University of Michigan, 1976); George H. Gallup, m e Gallup Poll: Public Opinion 1935-1971 (New York: Random House, 1972); and George H. Gallup, me Gallup Poll: Public Opinion, 1972-1977(Wilmington, Delaware: Scholarly Resources, 1978).

For most standard demographics only data from the GSSs are presented. Previous usages were too numerous and minor variations too common to merit the inclusion of these trends in the compendium. Trends for many of the demographics are available from such sources as Historical Statistics of the United States, Statistical Abstracts, or Social Indicators, 1976.

This compendium is also selective in that it does not cover all items contained in the GSSs. Counting the number of variables on a survey is largely an arbitrary exercise, depending on how filters are handled? whether summary scales are constructed from several questions or one question divided into several variables, and other considerations. The 1972-1978 General Social Survey cumulative data file comes up with 446 variables. To this total this compendium adds one additional item, COHORT or years of birth. It eliminates 152 variables: 82 items which were asked only a single time (38 one time experiments, 31 recent additions, 11 items dropped after their initial appearance, and 2 others), 39 items that were subparts of questions (e.g., specific Protestant demonination, and number of children expected in next five years), 18 Dictionary of Occupational_ Titles variables derived from occupation, 7 items subsumed under other variables (e.g., various income codes covered by generalized codes), 4 file documentation items (form, identification number, year, and sample code), and one miscellaneous item. This leaves 295 items ~- covered by the compendium. To this total 68 alternative codes for selected items (e.g., party identification or work statu~ are added. m is brings the number of variables covered up to 383. These additions create some redundancy (as even the supposedly 295 distinct items do by measuring such related matters as years of schooling and highest degree obtained and occupation and occupational prestige), but permit more detailed inspection of certain multi-category items.

Finally, since this compendium is a wholesale summary of many complex items and trends it cannot cover details, nuances, and other particularities of items, surveys, or trends. [Certain qualifications and elaborations are noted with the summary of trends. See also Appendix L, Changes in Question Wording, Response Categories, and Format in the GSS Cumulative Codebook.] For more comprehensive and intensive examination of particular trends one should either conduct a personal secondary analysis of the relevant data or, in instances where more complete analysis has already been conducted, consult the existing literature. (See Tom W. Smith, Annotated Bibliography-of Papers Using the General Social Surveys, Chicago: NORC, 1979), for a compilation of this body of material.)

Organization

The compendium consists of three sections: an introduction, a compilation of trends, and a subject index. The introduction gives the scope, organization, and summary of the selected trends. The main section lists trends in the distribution of the selected items. The subject index provides a topical guide to the items which should be particularly helpful to those uninitiated into the specialized language of GSS mnemonics.

In the main section, the items are arranged alphabetically by their GSS mnemonic with six exceptions. The variables evaluating various foreign countries (BRA2ILl, BRAZIL2, CANADAl, CANADA2, CHINAl, CHINA2, EGYPTl, EGYPT2, ENGLAND1, ENGLAND2, ISREALl, ISREAL2, RUSSIAl, and RUSSIA2) are grouped together under "Images of Countries," the variables ranking values for children (AMICABLE, CLEAN, CONSIDER, CONTROL, HONEST, INTEREST, JUDGEMENT, MANNERS, OBEYS, RESPONSI, ROLE, STUDIOUS, and SUCCESS) are grouped under "Qualities of Children," the variables ra~ing desirable occupational traits (JOBHOUR, JOBING, JOBMEANS, JOBPROMO, AND JOBSEC) are grouped under "Job Characteristics." !

Each entry has three parts: 1) question wording, 2) the selected trend proportions, and 3) the statistical analysis of the trend data. The question wording section gives the text of the item and defines how the responses categories

were grouped in order to calculate the proportion listed below. For example, for ABANY the proportion is the number of "Yes" responses (or cases) divided by the total number of responses ("Yes" responses plus "No" and "Don't know"). [In this and all other instances "no answer/n~ ascertained" and "not applicable/inappropriate," responses have been excluded from~he analysis.] The selected trend section gives the year that data were collected, the proportion as defined above, the total number of cases ("n"), and the study or survey name. Next it tries the linear model that all data points are random variations around a linear trend. The linear trend subsection has four parts: 1) "Weighted Regression," 2) "R Squared," 3) "Improvement~" and 4) "Fit." The "Weighted Regression" line gives the standard y = a + bx equation where "a" is the intercept and "b" is the slope or change per annum for the variable under analysis. [The intercepts are for the base year "O." As a result, they are arbitrary and interpretively meaningless constants.] The "R Squared" term measures the share of the variance in the trend variable that is explained by time. "Improvement" measures the difference between the fit of the "Constant" and the "Linear" models. "Fit" measures the goodness of fit between the best linear model and the observed data. The outcome of this test appears under the chi square and probability headings. Finally the statistical analysis section presents the best model description of the trend data based on the preceding fitting of constant and linear tests. There are five possible outcomes from these tests: 1) Constant, 2) Significant Linear Component, 3) Significant Linear trend, 4) No + Constant, not linear, and 5) Not Constant; can't decide model. The constant model is accepted when there is no significant variation from the constant or pooled proportion (i.e., the probability is .05 or greater). The significant linear component model is accepted when a) the constant model is rejected and b) the linear model is also rejected, but c) the linear model is a significant improvement over the constant model (i.e., the difference between the linear and constant models is significant at the .05 level). The significant linear trend model is accepted when a) the constant model is rejected and b) there is no significant variation from the linear fit (i.e., the probability is .05 or greater.) The ~fnot constant, not linear~model is accepted when a) the constant model is rejected, b) the linear model is rejected, and c) the improvement is not significant between the linear and constant models. The not constant; can't decide model outcome occurs when a) the constant model is rejected, b) the linear model is accepted, but c) the improvement is not significant. [This is a borderline occurance that infrequently appears.] The exact meaning of these models varies according to the exact configuration of the data and how one interprets them but can be described in general as follows:

```
1. Constant:    no significant change is occurring;
                no measurable trend appears

2. Significant linear component: significant change is occurring
                                  and there is a net direction to
                                  the trend; a trend appears but
                                  varies from a simple linear
                                  projection

3. Significant linear trend: change is occurring at a constant or
                             fixed rate per annum; the trend is
                             moving along a straight line

4. Not constant, not linear: significant change is occurring, but
                             it shows a complex pattern; the trend
                             does not follow a linear function
                             and has no simple direction
```

To illustrate this format further look at Figure 1. In part A the mnemonic ("ANOMIA5"), question wording, and

categorization of the proportion appears. In part B the four data points are presented. The time series covers the years 1973, 1974, 1976, and 1977. The proportion disagreeing with the item has ranged from a low of .376 in 1974 to a high of .433 in 1973. The case base ranges around 1,500 and all data come~ from the General Social Surveys (GSS). In part C the trend is statistically analyzed. In the Constant subsection the pooled estimate over the four ~-data points is .401 + .0126. The chi square and probability tests of significance show that the observed points differ significantly from the pooled estimate and therefore the constant model is rejected. The linear subsection reveals ehat the best linear estimate is that the item has been changing at a rate of .002 per annum from an intercept of 4.45. The association of the data points with time is a minimal .0196. The test for improvement finds that the best linear model is not signifi- cantly better than the constant model (i.e., the probability equals .618). The linear test shows that there is a significant amount of variation left after the best linear fit is tried (i.e., probability equals .001) so the linear model is also rejected. Since both the constant and the linear models were rejected, the model subsection reports that the trend in AMONIA5 over the four data points from 1973- ~ 1977 as ~not constant, not linear.~ Orjto put it another way~ the pro- portion disagreeing with the statement that the lot of the common man is getting worse varied significantly over this period, but the trend showed no net direction.

ANOMIA5

A In spite of what some people say, the lot (situation|condition) of the average man is getting worse, not better.

Proportion equals Disagree versus Agree and DK.

B

Year	proportion	N	Study
1973	0.4330	1502	GSS
1974	0.3760	1481	GSS
1976	0.38l0	1497	GSS
1977	0.4197	1525	GSS

C

TEST	ESTIMATE	CHI SQ	PROB
Constant	0.402 plus or minus 0.0126	14.9	0.002
Linear Trend			
Weighted regression.	Y = 4.45 - 0.0020(x)		
R Squared	0.0196		
Improvement		0.3	0.618
Fit		14.6	0.001

_ _

Model. Not constant, not linear

Fig. 1. Example of Format.

Summary of Trends

We will not try to substantively interpretate trends in this section. Trends on many particular GSS items are studied in papers cited in the Annotated Biblio~raphy. Some attempts to synthesize general trends in GSS items include James A. Davis, "Trends in NORC General Social Survey Items, 1972-1977," GSS Technical Report ~o. 9 (Chicago: NORC, 1978); James A. Davis, "Conservative Weather in a Liberalizing Climate: Change in Selected NORC General Social Survey Items, 1972- 1978," GSS Technical Report No. 13 (Chicago: NORC, 1979); and Tom W. Smith, "General Liberalism and Social Change in Post World War II America: A Summary of Trends," GSS Technical Report No.16 (Chicago: NORC, 1979).

What general trends and observations about the nature, extent, and direction of social change can be distilled out of these data depends in the first place on the amount and distribution of the actual trend data itself. It is extensive, but suffers from many limitations and inadequacies. For the 295 non-duplicates there are approximately 2100 data points. On the GSSs there were an average of 4.7 observations for each item. [Remember that all items with only single data point were excluded from the compendium.] Among the non-demographics (for which pre-GSS points were included), there were an average of 7.2 observations per item (4.4 GSS points and 2.9 non-GSS points). mere is, of course, a wide range in the actual number of pre-GSS points available for non-demographics from sixty-three items with no previous exact national usage to five items with twenty or more prior usages (CHLDIDEL-23, CLASS-21,CONBUS- 20,HAPPY-32, and USWAR-31).

Some of these Drior series stret~h b~k ~.~ f~r as the mid-thirtie~. but as the following figures show most come from only the last decade or so:

Time	Series Starting Durin~ Period
1930s	.025
1940s	.025
1950s	.151
1960s	.361
1970s	.437
	(238 - 57 demographics excluded)

The average starting time for the series was mid-1965, a fairly meaningless expression of central tendency. What both the figures on previous us~ages and time period covered point out is the large but very uneven amount of trend data that is contained in these series. One can look at aIl of the trends some of the time (the seventies) and some of the trends all of the time (or at least back to the beginning of national surveys}, but not all of the trends all of the

time. This imbalance in source data naturally hinders the analysis that one can carry out. We will deal with the problem of variable coverage by initially ignoring the problem and finally by "solving" it by restricting comparison to the data intensive GSS surveys. Looking at the final models for all non-duplicating series with three or more data points, [Since a two point trend series can only test out as constant or a significant linear trend these seventeen cases were deleted from the subsequent analysis.] we find that 26.0 percent were constant, 39.4 percent had linear components, 22.4 percent were linear trends, and 12.2 percent were non-constant, nonlinear. [One "can't decide model" cases was dropped from the analysis.] Adjusting for multi- stage sampling changed significantly the distribution of trend fits, however. [To allow for clustering the s~andard deviation were multiplied by 1.414.] With the adjustment 44.0 percent were constant, 29.6 percent had lie~ar components, 19.9 percent were linear trends, and 6.5 percent were non-constant, non-linear. By accepting more of the variation as sampling noise the adjusted models greatly simplify the types of models that can be fitted to the data. The proportion of change models (i.e., linear component, linear trend, and non-constant, non-linear) drops from 74 percent unadjusted to 56 percent when adjusted. Still even when the conservative cluster adjustment is made it appears that some change is occurring in a majority of items. Most of this change has some net direction to it, but it is usually more complex than simple linear change.

Of course this summary of trend models has been based on a hcdge- podge of time series covering from three to thirty-two data points and from 3 to 42 years and having restricted most demographics to GSS survey points. This can have great impact on the distribution of trend models since the types of models that fit series varies greatly by the length and density [We have not tried to seperate out the effects of series length and density but suspect that both factors independently lead to more complex models. The longer a series is, the more likely one is to detect as significant a small but steady trend. The more points in a series the more likely one is to detect 1) an episodic effect and 2) a cycical or non-linear trend.] of series and type of question. The impact of greater length and density can be seen by comparing the trend models for non- demographics on GSS surveys and on all surveys. On the GSS the unadjusted models were 43.4 percent constant, 20.7 percent linear component, 16.3 percent linear trend, and 20.7 percent non-constant, non-linear. Including the pre-GSS points shifted the models to 19.2 percent constant, 48.6 percent linear component, 18.3 percent linear trend, and 13.9 percent non-constant, non-linear. [Adjusted models show a similar pattern.] This would seem to indicate that if we had -- yearly readings for over thirty years for all of the items we would find even more change than presently appears or conversely with less points and shorter spans we would come up with appreciably less change than indicated above. [Other hypothese might explain this shift such as that change was greater in the period prior to the 1972-1978 span covered by the GSS or that house effects were causing the increase in variation between points. While neither is formally ruled out these are considered to be less likely hypotheses.]

Likewise the amount of change that is detected also varies by the question type. Four basic question types were distinguished: demographics which covered the basic background facts about a person's life, attitudes which were opinions on public issues and social values, personal evaluations which were self-rankings of psychological states, and behaviors which were reports of personal activities, memberships, or experiences. [The placing of particular items into one of these categorie~ was occasionally difficult as distinctions blurred. Three items (COOP, COMPREND, and ETHNUM) were left in a miscellaneous category and about a half dozen other items were fairly arbritarily assigned to categories.] This breakdown pretty much follows traditional dis- tinctions in the classification of survey items with the exception that personal evaluations are usually grouped with attitudes.

Attitudes show the most change models (66.2 percent) followed by demographics (48.4 percent), behaviors (35.0 percent), and personal evaluations (16.3 percent) (Table 1). Attitude items led in the number of trends that had linear components and non-constant, non-linear trends.

TABLE 1

FINAL MODELS BY QUESTION TYPE

(GSS only: unadjusted)

Trend Model	Question Type				
	Demographics	Attitudes	Evaluations	Behaviors	Total
Constant	51.6%	33.8%	83.3%	65.0%	45.9%
Linear Component.	9.4	25.7	8.3	5.0	17.3
Linear Trend	31.3	14.7	0.0	25.0	20.0
Non-Constant, Non-Linear	7.8	25.7	8.3	5.0	16.9
	(64)	(137)	(12)	(40)	(256)

Includes three miscellaneous items not categorized in table.

Includes one "can't decide model" case that is excluded from the percentaging.

This indicates that attitude items not only showed more change, but also were more likely to either follow more complex trends or to bounce around. Demographics were next most likely to show change, but were much more likely to have simple linear trends than attitude items were. Behavioral items showed less change and followed the pattern of the demographics in havin~ moY~ly linear chan~e model~. Per~nql ~v~luati~n~ ~r~ h~e;~ constant and with a case base of 12~ there is insufficient data for more detailed analysis.

The four categories of items also varied in the amount of linear change. Among items testing out as linear components or linear trends on the GSS attitude items had an average slope of 1.4 percent per annum, demographics averaged 1.15 percent, behaviors 1.29 percent, and the only personal evaluation with a linear component had a slope of 0.59 percent. It appears that attitudes not only were more likely to show change, but among those with linear change attitudes had the largest annual rate of change.

Clearly the amount and type of change measured would vary widely according to what mixture of question types is placed in the trend salad. Because of this it is impossible to come up with any meaningful simple summary of the amount and type of change that has been occurring. Still some generalizations are possible. First, even allowing for much artifact it is evident that much change has occurred. Even for the clustering adjusted figures for all points, the items showed non-constant trends for 56 percent of the time series. Second, the change that does occur usually has some net direction to it, but it is often more complex than a simple linear trend. m ird, linear change usually occurs at something less than a banner headline pace. Attitudes, demographics, and behaviors that show linear change averaged only 1.3 percent on the GSSs ~ 1.1 percent for all data points. [With more data points over a longer average time span it was possible for more small trends to show up as significant when all points were used rather than just GSS points. Calculating in all the constant trends would, of course, reduce these average annual rates of change by a notable margin. Using the adjusted models would on the other hand increase the rate of change since it reclassifies many of the weaker linear trends as constant.] The greatest rate of change on the GSS surveys was only 3.2 percent per annum (an inflation triggered increase in respondent's income) and on all surveys only 6 percent of all linear change was above 2.5 percent per annum. [Some non-linear trends were, of course, larger. The largest annual shift on the GSSs was a drop of 19.9 percent in confidence in organized religion between 1974 and 1975 and the largest standard deviation from a constant trend was .0729 for confidence in the executive branch of the federal government.]

Fourth, change varies by the type of question. If we are willing to accept the GSS items as typical representatives for each of the four types of questions distinguished above, it appears that attitudes show the most change and the most complex change models. Demographics are more stable than attitudes and are more likely to show simple linear trends than linear components or non-linear trends. Behaviors and personal evaluations are even more constant than either attitudes or demographics. [A listing of the twenty most changable items (largest standard deviations from their pooled average) on the GSSs confirmed this picture Seventeen were attitudes, two were demographics, one was a behavior, and none were personal evaluations. The expected distribution would have been 10.8 attitudes, 5.1 demographics, 3.2 behaviors, and 0.9 personal evaluation.] Of this pattern perhaps the most surprising features are that demographics are not quite as rock solid as sometimes pictured and that perhaps the most subjective of all question types, personal evaluations, are the most stable.

In sum, the trend data contained in this compendium shows a satis- factorily complicated pattern of stability, linear and more complex changes, and bounce. Many items show major changes while many others show great consistancy. Although far from being comprehen~ive and by necessity sacrificing depth for breath, it provides a useful summary of many recent trends and will hopefully serve as a helpful reference source for studying social change.