SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Chapter 1
Social Issues and the Study of American Institutions

As an introduction to the study of social issues, Skolnick and Currie's history of changes in the study of American social institutions (2000:1-13) provides a useful background for student projects based on this data set. These broad trends in the study of social problems and social changes in America both reflect and affect our basic assumptions about society and the way social issues and social problems are perceived by Americans in general and by social science as it has analyzed them.

At the end of the Nineteenth Century, American values and institutions emphasized individual hard work, thrift, and personal discipline. The challenge to American society was to maintain these distinctive American values despite industrialization and urbanization, accompanied by immigration from other parts of the world. Defective moral characteristics in individuals were considered to be the cause of social problems. Social scientists, politicians, and social reformers sought ways to change these people into individuals who could compete and succeed in American society. The solutions they proposed emphasized social control such as prisons and mental institution but included a few social welfare programs. Social scientists concentrated on the scientific study of the society to preserve its optimal functioning. Their personal or professional values as social scientists or as social reformers were not to be involved. Social problems were considered to be signs of problems in a particular segment of society. After objective study, social scientists would make recommendations for change to government or business. The ideals and values of American society were accepted without question, especially the competitive, capitalistic economic system characterized by private property and individual competition.

World War I, the Depression, and World War II interrupted this study of American life, but the end of World War II brought political optimism and economic affluence for many Americans. Often this was the first real chance for people to meet economic and personal goals. The expanding economy produced jobs that paid men well enough that they could support families. The 'baby boom" was really a nuclear family boom as Americans married at higher rates and younger ages and had more babies. Communism was considered to be the most serious threat to American culture and economic-political life. Any social changes were left to the expertise of the political and military institutions and were not social science concerns. Social scientists

supported existing social American institutions, considering a strong national defense and effective counterespionage intelligence to be necessary and desirable.

During the 1960s American perspectives shifted as the country became more aware of disadvantaged people both in our general affluence here and in so-called underdeveloped areas elsewhere. At first the response was to extend American technological and political resources to the less fortunate, especially encouraging social changes to help people help themselves toward democracy, development, and modernization. American society was considered to be the ideal economic and political system. (The only criticism was that not enough of the world, or even of our own people, benefited from it.) The early 60s were optimistic that this could be done, and social science concerned itself with identifying glitches in the system, focusing on particular social problems as the deviant behavior of individuals or social disorganization in segments of society. It was assumed that scientists could recommend appropriate changes. Social analysis was assumed to be politically neutral. If the society operated less efficiently than it might, specific problems would be analyzed and then referred to the appropriate social institutions—education, political, military--for adjustments. During the late 1960s and early 1970s, the federal government began a variety of social programs to bring American reality closer to the ideal. Programs and legislation included the War on Poverty, the Civil Rights Act, Medicare expanding the Social Security system, the proposed Equal Rights Amendment, Title IX of the Education Amendment to the Civil Rights Act.

However, later in the 1970s, economic problems, persistent poverty, racial and ethnic cleavages, urban disorganization, increasing crime and violence led Americans, including many social scientists, to more pessimistic conclusions. Some concluded that the government had tried to do too much for people. Maybe the programs were too generous and had negative consequences in the long term. Theories of racial inferiority and cultural inadequacy revived. Harsher sentences for of those convicted of crimes were mandated, and some states reinstated the death penalty. Communities spent more money for prisons and less for education. By mid 1990s, welfare reform legislation was designed to force the poor to work and limited the time for benefits for their families. Public opinion on social issues showed cleavages within the public that were taken more seriously, not only by public officials concerned with reelection, but also by social scientists and the general public. Public debate and controversy increased and became almost a phenomenon its own right. Since problems such as poverty, crime, school failure continued despite government programs and social policies, the conclusion that these programs had failed or even contributed to ongoing problems seemed plausible. The idea that the disadvantages stemmed from deficiencies in individuals, families, communities, and/or subcultures reappeared.

By the beginning of the twenty-first century, American thinking about social issues seemed to have come full circle and now blamed school failure, poverty, delinquency, and welfare dependence on individuals or subcultures. At the same time, the gap between the have and have-nots increased, and the American economic system had been transformed by global economic competition and new workplace technology. There seemed to be no consensus on solving the problems related to the increasingly complex and rapid changes that affected many American institutions.

REFERENCES AND SUGGESTED READING

Social Issues

- Skolnick, Jerome H. and Elliot Currie. 2000. *Crisis in American Institutions* Eleventh Edition. Boston: Allyn and Bacon

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Chapter 2
Survey Research Design and Quantitative Methods of Analysis for Cross-Sectional Data

Almost everyone has experience with surveys. Market surveys ask respondents whether they recognize products and their feelings about them. Political polls ask questions about candidates for political office or opinions related to political and social issues. Needs assessments use surveys that identify the needs of groups. Evaluations often use surveys to assess the extent to which programs achieve their goals.

Survey research is a method of collecting information by asking questions. Sometimes interviews are done face-to-face with people at home, in school, or at work. Other times questions are sent in the mail for people to answer and mail back. Increasingly, surveys are conducted by telephone and over the internet.

SAMPLE SURVEYS

Although we want to have information on all people, it is usually too expensive and time consuming to question everyone. So we select only some of these individuals and question them. It is important to select these people in ways that make it likely that they represent the larger group.

The **population** is all the objects in which we are interested. Often populations consist of individuals. For example, a population might consist of all adults living in California. But it may also be geographical areas such as all cities with populations of 100,000 or more. Or we may be interested in all households in a particular area. A **sample** is the subset of the population involved in a study. In other words, a sample is part of the population. The process of selecting the sample is called **sampling**. The idea of sampling is to select part of the population to represent the entire population.

The United States Census is a good example of sampling. The census tries to enumerate all residents every ten years with a short questionnaire. In 2000, approximately one out every six households was given a longer questionnaire. Information from

this sample (i.e., every sixth household) was used to make inferences about the population. In the future, approximately 250,000 households will be sampled every month. Political polls also use samples. To find out how potential voters feel about a particular political race, pollsters select a sample of potential voters. This module uses opinions from a sample of adults (18+) living in the United States collected at several points in time.

Since a survey can be no better than the quality of the sample, it is essential to understand the basic principles of sampling. There are two types of sampling-probability and nonprobability. A **probability sample** is one in which each individual in the population has a known, nonzero, chance of being selected in the sample. The most basic type is the **simple random sample**. In a simple random sample, every individual (and every combination of individuals) has the same chance of being selected in the sample. This is the equivalent of writing each person's name on a piece of paper, putting them in plastic balls, putting all the balls in a big bowl, mixing the balls thoroughly, and selecting some predetermined number of balls from the bowl. This would produce a simple random sample.

The simple random sample assumes that we can list all the individuals in the population, but often this is impossible. If our population were all the households or residents of California, there would be no list of the households or residents available, and it would be very expensive and time consuming to construct one. In this type of situation, a **multistage cluster sample** would be used. The idea is very simple. If we wanted to draw a sample of all residents of California, we might start by dividing California into large geographical areas such as counties and selecting a sample of these counties. Our sample of counties could then be divided into smaller geographical areas such as blocks and a sample of blocks would be selected. We could then construct a list of all households for only those blocks in the sample. Finally, we would go to these households and randomly select one member of each household for our sample. Once the household and the member of that household have been selected, substitution would not be allowed. This often means that we must call back many times, but this is the price we must pay for a good sample.

Telephone samples often use a technique called **random-digit dialing**. With random-digit dialing, phone numbers are dialed randomly within working exchanges. Numbers are selected in such a way that all areas have the proper proportional chance of being selected in the sample. Random-digit dialing makes it possible to include numbers that are not listed in the telephone

directory and households that have moved into an area so recently that they are not included in the current telephone directory.

A **nonprobability sample** is one in which each individual in the population does not have a known chance of selection in the sample. There are several types of nonprobability samples. For example, magazines often include questionnaires for readers to fill out and return. This is a **volunteer sample** since respondents self-select themselves into the sample (i.e., they volunteer to be in the sample). Another type of nonprobability sample is a **quota sample**. Survey researchers may assign quotas to interviewers. For example, interviewers might be told that half of their respondents must be female and the other half male. This is a quota on sex. We could also have quotas on several variables (e.g., sex and race) simultaneously.

Probability samples are preferable to nonprobability samples. First, they avoid the dangers of what survey researchers call "systematic selection biases" which are inherent in nonprobability samples. For example, in a volunteer sample, particular types of persons might be more likely to volunteer. Perhaps highly-educated individuals are more likely to volunteer to be in the sample and this would produce a systematic selection bias in favor of the highly educated. In a probability sample, the selection of the actual cases in the sample is left to chance. Second, in a probability sample we are able to estimate the amount of sampling error (our next concept to discuss).

We would like our sample to give us a perfectly accurate picture of the population. However, this is unrealistic. Assume that the population is all employees of a large corporation, and we want to estimate the percent of employees in the population that is satisfied with their jobs. We select a simple random sample of 500 employees and ask the individuals in the sample how satisfied they are with their jobs. We discover that 75 percent of the employees in our sample are satisfied. Can we assume that 75 percent of the population is satisfied? That would be asking too much. Why would we expect one sample of 500 to give us a perfect representation of the population? We could take several different samples of 500 employees and the percent satisfied from each sample would vary from sample to sample. There will be a certain amount of error as a result of selecting a sample from the population. We refer to this as **sampling error**. Sampling error can be estimated in a probability sample, but not in a nonprobability sample.

It would be wrong to assume that the only reason our sample estimate is different from the true population value is because

of sampling error. There are many other sources of error called **nonsampling error**. Nonsampling error would include such things as the effects of biased questions, the tendency of respondents to systematically underestimate such things as age, the exclusion of certain types of people from the sample (e.g., those without phones, those without permanent addresses, those we are never able to contact, those who refuse to answer our questions), or the tendency of some respondents to systematically agree to statements regardless of the content of the statements. In some studies, the amount of nonsampling error might be far greater than the amount of sampling error. Notice that sampling error is random in nature, while nonsampling error may be nonrandom producing systematic biases. We can estimate the amount of sampling error (assuming probability sampling), but it is much more difficult to estimate nonsampling error. We can never eliminate sampling error entirely, and it is unrealistic to expect that we could ever eliminate nonsampling error. It is good research practice to be diligent in seeking out sources of nonsampling error and trying to minimize them.

DATA ANALYSIS: Examining Variables One at a Time (Univariate Analysis)

The rest of this chapter will deal with the analysis of survey **data**. Data analysis involves looking at variables or "things" that vary or change. A **variable** is a characteristic of the individual (assuming we are studying individuals). The answer to each question on the survey forms a variable. For example, sex is a variable-some individuals in the sample are male and some are female. Age is a variable; individuals vary in their ages.

Looking at variables one at a time is called **univariate analysis**. This is the usual starting point in analyzing survey data. There are several reasons to look at variables one at a time. First, we want to describe the data. How many of our sample are men and how many are women? How many are African-Americans and how many are white? What is the distribution by age? How many say they are going to vote for Candidate A and how many for Candidate B? How many respondents agree and how many disagree with a statement describing a particular opinion?

Another reason we might want to look at variables one at a time involves recoding. **Recoding** is the process of combining categories within a variable. Consider age, for example. In the data set used in this module, age varies from 18 to 89, but we

would want to use fewer categories in our analysis, so we might combine age into age 18 to 29, 30 to 49, and 50 and over. We might want to combine African Americans with the other races to classify race into only two categories-white and nonwhite. Recoding is used to reduce the number of categories in the variable (e.g., age) or to combine categories so that you can make particular types of comparisons (e.g., white versus nonwhite).

The frequency distribution is one of the basic tools for looking at variables one at a time. A **frequency distribution** is a set of categories and the number of cases in each category. **Percent distributions** show the percentage in each category. Table 2.1 shows frequency and percent distributions for two hypothetical variables-one for sex and one for willingness to vote for a woman candidate. Begin by looking at the frequency distribution for sex. There are three columns in this table. The first column specifies the categories-male and female. The second column tells us how many cases there are in each category, and the third column converts these frequencies into percents.

Table 2.1 -- Frequency and Percent Distributions for Sex and Willingness to Vote for a Woman Candidate (Hypothetical Data)

Sex			**Voting Preference**			
Category	Freq.	Percent	Category	Freq.	Percent	Valid Percent
Male	380	40.0	Willing to Vote for a Woman	460	48.4	51.1
Female	570	60.0	Not Willing to Vote for a Woman	440	46.3	48.9
Total	950	100.0	Refused	50	5.3	Missing
			Total	950	100.0	100.0

In this hypothetical example, there are 380 males and 570 females or 40 percent male and 60 percent female. There are a total of 950 cases. Since we know the sex for each case, there

are no **missing data** (i.e., no cases where we do not know the proper category). Look at the frequency distribution for voting preference in Table 2.1. How many say they are willing to vote for a woman candidate and how many are unwilling? (Answer: 460 willing and 440 not willing) How many refused to answer the question? (Answer: 50) What percent say they are willing to vote for a woman, what percent are not, and what percent refused to answer? (Answer: 48.4 percent willing to vote for a woman, 46.3 percent not willing, and 5.3 percent refused to tell us.) The 50 respondents who didn't want to answer the question are called missing data because we don't know which category into which to place them, so we create a new category (i.e., refused) for them. Since we don't know where they should go, we would want a percentage distribution considering only the 900 respondents who answered the question. We can determine this easily by taking the 50 cases with missing information out of the base (i.e., the denominator of the fraction) and recomputing the percentages. The fourth column in the frequency distribution (labeled "valid percent") gives us this information. Approximately 51 percent of those who answered the question were willing to vote for a woman and approximately 49 percent were not.

With these data we will use frequency distributions to describe variables one at a time. There are other ways to describe single variables. The mean, median, and mode are averages that may be used to describe the central tendency of a distribution. The range and standard deviation are measures of the amount of variability or dispersion of a distribution. (We will not be using measures of central tendency or variability in this module.)

Exploring the Relationship Between Two Variables (Bivariate Analysis)

Usually we want to do more than simply describe variables one at a time. We may want to analyze the relationship between variables. Morris Rosenberg (1968:2) suggests that there are three types of relationships: "(1) neither variable may influence one another .... (2) both variables may influence one another ... (3) one of the variables may influence the other." We will focus on the third of these types which Rosenberg calls "asymmetrical relationships." In this type of relationship, one of the variables (the **independent variable**) is assumed to be the cause and the other variable (the **dependent variable**) is assumed to be the effect. In other words, the independent variable is the variable that influences the dependent variable.

For example, researchers think that smoking causes lung cancer. The statement that specifies the relationship between two variables is called a **hypothesis** (see Hoover and Donovan, 2004, for a more extended discussion of hypotheses). In this hypothesis, the independent variable is smoking (or more precisely, the amount one smokes) and the dependent variable is lung cancer. Consider another example. Political analysts think that income influences voting decisions, that rich people vote differently from poor people. In this hypothesis, income would be the independent variable and voting would be the dependent variable.

In order to demonstrate that a **causal relationship** exists between two variables, we must meet three criteria: (1) there must be a statistical relationship between the two variables, (2) we must be able to demonstrate which one of the variables influences the other, and (3) we must be able to show that there is no other alternative explanation for the relationship. As you can imagine, it is impossible to show that there is no other alternative explanation for a relationship. For this reason, we can show that one variable does not influence another variable, but we cannot prove that it does. We can only show that a causal relationship is plausible or credible. In this section, we will focus on the first two criteria and leave this third criterion to the next section.

In the previous section we looked at the frequency distributions for sex and voting preference. All we can say from these two distributions is that the sample is 40 percent men and 60 percent women and that slightly more than half of the respondents said they would be willing to vote for a woman, and slightly less than half are not willing to. We cannot say anything about the relationship between sex and voting preference. In order to determine if men or women are more likely to be willing to vote for a woman candidate, we must move from univariate to bivariate analysis.

A **crosstabulation** (or **contingency table**) is the basic tool used to explore the relationship between two variables. Table 2.2 is the crosstabulation of sex and voting preference. In the lower right-hand corner is the total number of cases in this table (900). Notice that this is not the number of cases in the sample. There were originally 950 cases in this sample, but any case that had missing information on either or both of the two variables in the table has been excluded from the table. Be sure to check how many cases have been excluded from your table and to indicate this figure in your report. Also be sure that you understand why these cases have been excluded. The figures in the lower margin and right-hand margin of the table are called

the marginal distributions. They are simply the frequency distributions for the two variables in the table. Here, there are 360 males and 540 females (the marginal distribution for the column variable-sex) and 460 people who are willing to vote for a woman candidate and 440 who are not (the marginal distribution for the row variable-voting preference). The other figures in the table are the cell frequencies. Since there are two columns and two rows in this table (sometimes called a 2 x 2 table), there are four cells. The numbers in these cells tell us how many cases fall into each combination of categories of the two variables. This sounds complicated, but it isn't. For example, 158 males are willing to vote for a woman and 302 females are willing to vote for a woman.

Table 2.2 -- Crosstabulation of Sex and Voting Preference (Frequencies)

	Sex		
Voting Preference	Male	Female	Total
Willing to Vote for a Woman	158	302	460
Not Willing to Vote for a Woman	202	238	440
Total	360	540	900

We could make comparisons rather easily if we had an equal number of women and men. Since these numbers are not equal, we must use percentages to help us make the comparisons. Since percentages convert everything to a common base of 100, the percent distribution shows us what the table would look like if there were an equal number of men and women.

Before we percentage Table 2.2, we must decide which of these two variables is the independent and which is the dependent variable. Remember that the independent variable is the variable we think might be the influencing factor. The independent variable is hypothesized to be the cause, and the dependent variable is the effect. Another way to express this is to say that the dependent variable is the one we want to explain. Since we think that sex influences willingness to vote for a woman candidate, sex would be the independent variable.

Once we have decided which is the independent variable, we are ready to percentage the table. Notice that percentages can be computed in different ways. In Table 2.3, the percentages have been computed so that they sum down to 100. These are called **column percents**. If they sum across to 100, they are called **row percents**. If the independent variable is the column variable, then we want the percents to sum down to 100 (i.e., we want the column percents). If the independent variable is the row variable, we want the percents to sum across to 100 (i.e., we want the row percents). This is a simple, but very important, rule to remember. We'll call this our **rule for computing percents**. Although we often see the independent variable as the column variable so the table sums down to 100 percent, it really doesn't matter whether the independent variable is the column or the row variable. In this module, we will put the independent variable as the column variable. Many others (but not everyone) use this convention. Please do this when you write your report.

Table 2.3 -- Voting Preference by Sex (Percents)

Voting Preference	Male	Female	Total
Willing to Vote for a Woman	43.9	55.9	51.1
Not Willing to Vote for a Woman	56.1	44.1	48.9
Total Percent	100.0	100.0	100.0
(Total Frequency)	(360)	(540)	(900)

Now we are ready to interpret this table. Interpreting a table means to explain what the table is saying about the relationship between the two variables. First, we can look at each category of the independent variable separately to describe the data and then we compare them to each other. Since the percents sum down to 100 percent, we compare across. The **rule for interpreting percents** is to compare in the direction opposite to the way the percents sum to 100. So, if the percents sum down to 100, we compare across, and if the percents sum across to 100, compare down. If the independent variable is the column variable, the percents will **always** sum down to 100. We can look at each category of the independent variable separately to describe the data and then compare them to each other-describe down and then compare across. In Table 2.3, row one

shows the percent of males and the percent of females who are willing to vote for a woman candidate--43.9 percent of males are willing to vote for a woman, while 55.9 percent of the females are. This is a difference of 12 percentage points. Somewhat more females than males are willing to vote for a woman. The second row shows the percent of males and females who are not willing to vote for a woman. Since there are only two rows, the second row will be the complement (or the reverse) of the first row. It shows that males are somewhat more likely to be unwilling to vote for a woman candidate (a difference of 12 percentage points in the opposite direction).

When we observe a difference, we must also decide whether it is significant. There are two different meanings for significance-statistical significance and substantive significance. **Statistical significance** considers whether the difference is great enough that it is probably not due to chance factors. **Substantive significance** considers whether a difference is large enough to be important. With a very large sample, a very small difference is often statistically significant, but that difference may be so small that we decide it isn't substantively significant (i.e., it's so small that we decide it doesn't mean very much). We're going to focus on statistical significance, but remember that even if a difference is statistically significant, you must also decide if it is substantively significant.

Let's discuss this idea of statistical significance. If our population is all adults in the United States, we want to know if there is a relationship between sex and voting preference in this population. All we have is information about a sample from the population. We use the sample information to make an inference about the population. This is called **statistical inference**. We know that our sample is not a perfect representation of our population because of **sampling error**. Therefore, we would not expect the relationship we see in our sample to be exactly the same as the relationship in the population.

Suppose we want to know whether there is a relationship between sex and voting preference in the population. It is impossible to prove this directly, so we have to demonstrate it indirectly. We set up a hypothesis (called the **null hypothesis**) that says that sex and voting preference are not related to each other in the population. This basically says that any difference we see is likely to be the result of random variation. If the difference is large enough that it is not likely to be due to chance, we can reject this null hypothesis of only random differences. Then the hypothesis that they are related (called

the **alternative or research hypothesis**) will be more credible.

Table 2.4 -- Computation of Chi Square Statistic

Column 1	Column 2	Column 3	Column 4	Column 5
f_o	f_e	$(f_o - f_e)$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
158	184	-26	676	3.67
202	176	26	676	3.84
302	276	26	676	2.45
238	264	-26	676	2.56

12.52 = chi square

In the first column of Table 2.4, we have listed the four cell frequencies from the crosstabulation of sex and voting preference. We'll call these the **observed frequencies** (f_o) because they are what we observe from our table. In the second column, we have listed the frequencies we would expect if, in fact, there is no relationship between sex and voting preference in the population. These are called the **expected frequencies** (f_e). We'll briefly explain how these expected frequencies are obtained. Notice from Table 2.3 that 51.1 percent of the sample were willing to vote for a woman candidate, while 48.9 percent were not. If sex and voting preference are independent (i.e., not related), we should find the same percentages for males and females. In other words, 48.9 percent (or 176) of the males and 48.9 percent (or 264) of the females would be unwilling to vote for a woman candidate. (This explanation is adapted from Norusis, 2005.) Now, we want to compare these two sets of frequencies to see if the observed frequencies are really like the expected frequencies. All we do is to subtract the expected from the observed frequencies (column three). We are interested in the sum of these differences for all cells in the table. Since they always sum to zero, we square the differences (column four) to get positive numbers. Finally, we divide this squared difference by the expected frequency (column five). (Don't worry about why we do this. The reasons are technical and don't add to your understanding.) The sum of column five (12.52) is called the **chi square statistic**. If the observed and the expected frequencies are identical (no difference), chi square will be zero. The greater

the difference between the observed and expected frequencies, the larger the chi square.

If we get a large chi square, we are willing to reject the null hypothesis. How large does the chi square have to be? We reject the null hypothesis of no relationship between the two variables when the probability of getting a chi square this large or larger by chance is so small that the null hypothesis is very unlikely to be true. That is, if a chi square this large would rarely occur by chance (usually less than once in a hundred or less than five times in a hundred). In this example, the probability of getting a chi square as large as 12.52 or larger by chance is less than one in a thousand. This is so unlikely that we reject the null hypothesis, and we conclude that the alternative hypothesis (i.e., there is a relationship between sex and voting preference) is credible (not that it is necessarily true, but that it is credible). There is always a small chance that the null hypothesis is true even when we decide to reject it. In other words, we can never be sure that it is false. We can only conclude that there is little chance that it is true.

Just because we have concluded that there is a relationship between sex and voting preference does not mean that it is a strong relationship. It might be a moderate or even a weak relationship. There are many statistics that measure the strength of the relationship between two variables. Chi square is not a measure of the strength of the relationship. It just helps us decide if there is a basis for saying a relationship exists regardless of its strength. **Measures of association** estimate the strength of the relationship and are often used with chi square. (See Appendix D for a discussion of how to compute the two measures of association discussed below.)

Cramer's V is a measure of association appropriate when one or both of the variables consists of unordered categories. For example, race (white, African American, other) or religion (Protestant, Catholic, Jewish, other, none) are variables with unordered categories. Cramer's V is a measure based on chi square. It ranges from zero to one. The closer to zero, the weaker the relationship; the closer to one, the stronger the relationship.

Gamma (sometimes referred to as Goodman and Kruskal's Gamma) is a measure of association appropriate when both of the variables consist of ordered categories. For example, if respondents answer that they strongly agree, agree, disagree, or strongly disagree with a statement, their responses are ordered. Similarly, if we group age into categories such as under 30, 30 to 49, and 50 and over, these categories would be

ordered. Ordered categories can logically be arranged in only two ways-low to high or high to low. Gamma ranges from zero to one, but can be positive or negative. The sign of Gamma is arbitrary since it depends on the way the categories are arranged. So it's best to ignore the sign and focus on the numerical value. You can use the percentages to decide on the direction (i.e., positive or negative) of the relationship. Like V, the closer to zero, the weaker the relationship and the closer to one, the stronger the relationship.

Choosing whether to use Cramer's V or Gamma depends on whether the categories of the variable are ordered or unordered. However, dichotomies (variables consisting of only two categories) may be treated as if they are ordered. For example, sex is a **dichotomy** consisting of the categories male and female. There are only two possible ways to order sex-male, female and female, male. Or, race may be classified into two categories-white and nonwhite. We can treat dichotomies as if they consisted of ordered categories because they can be ordered in only two ways. In other words, when one of the variables is a dichotomy, treat this variable as if it were ordinal and use gamma. This is important when choosing an appropriate measure of association.

In this chapter we have described how surveys are done and how we analyze the relationship between two variables. In the next chapter we will explore how to introduce additional variables into the analysis.

REFERENCES AND SUGGESTED READING

Methods of Social Research

- Earl Babbie, 2007, *The Practice of Social Research* (11th edition), Thomson Wadsworth.
- Kenneth R. Hoover and Todd Donovan, 2004, *The Elements of Social Scientific Thinking* (8th Ed.). Thomson Wadsworth.

Interviewing

- Gorden, Raymond L. 1987. *Interviewing: Strategy, Techniques and Tactics*, Dorsey.

Survey Research and Sampling

- Earl Babbie, 2007, *The Practice of Social Research* (11th edition), Thomson Wadsworth.
- Arlene Fink, Linda Bourque, Eve P. Fieldler, Sabrina Mertens Oishi, Mark S. Litwin, 2003, The Survey Kit, Sage.
- Don A. Dillman, 2000, Mail and Internet Surveys – The Tailored Design Method (2nd edition), John Wiley & Sons.

Statistical Analysis

- Herman J. Loether and Donald G. McTavish, 1993, *Descriptive and Inferential Statistics An Introduction* (4th edition), Allyn and Bacon.
- J. Richard Kendrick, Jr., 2000, *Social Statistics*, Mayfield.
- Marija J. Norusis, 2005, *SPSS 13.0 Guide to Data Analysis*, Prentice Hall.

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Chapter 3
Introducing a Control Variable (Multivariate Analysis)

Human behavior is usually too complicated to be studied with only two variables. Often we will want to consider sets of three or more variables (called **multivariate analysis**). We will want to consider three or more variables when we have discovered a relationship between two variables and want to find out (1) if this relationship might be due to some other factor, (2) how or why these variables are related, or (3) if the relationship is the same for different types of individuals.

In each situation, we identify a third variable that we want to consider. This is called the **control** or the **test variable**. (Although it is possible to use several control variables simultaneously, we will limit ourselves to one control variable at a time.) To introduce a third variable, we identify the control variable and separate the cases in our sample by the categories of the control variable. For example, if the control variable is age divided into these two categories--younger and older, we would separate the cases into two groups. One group would consist of individuals who are younger and the other group would be those who are older. We would then obtain the crosstabulation of the independent and dependent variables for each of these age groups. Since there are two categories in this control variable, we obtain two **partial tables**, each containing part of the original sample. (If there were three categories in our control variable, for example, young, middle aged, and old, we would have three partial tables.)

The process of using a control variable in the analysis is called **elaboration** and was developed at Columbia University by Paul Lazarsfeld, Patricia Kendall, and their associates. There are several different types of outcomes to the elaboration process. We will discuss each briefly.

Table 2.3 showed that females were more likely than males to say they were willing to vote for a woman. Let's introduce a control variable and see what happens. In this example we are going to use age as the control variable.

Table 3.1 is the three-variable table with voting preference as the dependent variable, sex as the independent variable, and age as the control variable. When we look at the older respondents (the left-hand partial table), we discover that this partial table is very similar to the original two-variable table (Table 2.3). The same is true for the younger respondents (the right-hand partial table). Each partial table is very similar to the original two-variable table. This is often referred to as **replication** because the partial tables repeat the original two-variable table (see Babbie, 2007: 437). It is not necessary that they be identical; just that each partial table be basically the same as the original two-variable table. Our conclusion is that age is not affecting the relationship between sex and voting preference. In other words, the difference between males and females in voting preference is not due to age.

Table 3.1 -- Voting Preference by Sex Controlling for Age

	Older			**Younger**		
	Male %	Female %	Total %	Male %	Female %	Total %
Voting Preference						
Willing to Vote for a Woman	43.8	56.1	49.0	44.2	55.8	52.9
Not Willing to Vote for a Woman	56.2	43.9	51.0	55.8	44.2	47.1
	100.0	100.0	100.0	100.0	100.0	100.0
	(240)	(180)	(420)	(120)	(360)	(480)

Since this is a hypothetical example, imagine a different outcome. Suppose we introduce age as a control variable and instead of getting Table 3.1, we get Table 3.2. How do these two tables differ? In Table 3.2, the percentage difference between males and females has disappeared in both of the partial tables. This is called **explanation** because the control variable, age, has explained away the original relationship between sex and voting preference. (We often say that the relationship between the two variables is **spurious**, not causal.) When age is held constant, the difference between males and

females disappears. The difference in the relationship does not have to disappear entirely, only be reduced substantially in each of the partial tables. This can only occur when there is a relationship between the control variable (age) and each of the other two variables (sex and voting preference).

Next, we are interested in how or why the two variables are related. Suppose females are more likely than males to vote for a woman and that this difference cannot be explained away by age or by any other variable we have considered. We need to think about why there might be such a difference in the preferences of males and females. Perhaps females are more

Table 3.2 -- Voting Preference by Sex Controlling for Age

	Older			**Younger**		
	Male %	Female %	Total %	Male %	Female %	Total %
Voting Preference						
Willing to Vote for a Woman	32.9	33.9	33.3	65.8	66.9	66.7
Not Willing to Vote for a Woman	67.1	66.1	66.7	34.2	33.1	33.3
	100.0	100.0	100.0	100.0	100.0	100.0
	(240)	(180)	(420)	(120)	(360)	(480)

liberal than males, and liberals are more likely to say they would vote for a woman. So we introduce liberalism/conservatism as a control variable in our analysis. If females are more likely to support a woman because they are more liberal, then the difference between the preferences of men and women should disappear or be substantially reduced when liberalism/conservatism is held constant. This process is called **interpretation** because we are interpreting how one variable is related to another variable. Table 3.3 shows what we would expect to find if females supported the woman because they were more liberal. Notice that in both partial tables, the differences in the percentages between men and women has disappeared. (It is not necessary that it disappears entirely, but only that it is substantially reduced in each of the partial tables.)

The crucial difference between explanation and interpretation is that in interpretation the control variable is an intervening variable. The independent variable is causally prior to the control variable and the control variable is causally prior to the dependent variable. The control variable provides the link between the independent and dependent variables. It helps us interpret how the independent variable leads to the dependent variable.

Table 3.3 -- Voting Preference by Sex Controlling for Liberalism/Conservatism

	Conservative			**Liberal**		
	Male %	Female %	Total %	Male %	Female %	Total %
Voting Preference						
Willing to Vote for a Woman	32.9	33.9	33.3	65.8	66.9	66.7
Not Willing to Vote for a Woman	67.1	66.1	66.7	34.2	33.1	33.3
	100.0	100.0	100.0	100.0	100.0	100.0
	(240)	(180)	(420)	(120)	(360)	(480)

Finally, let's focus on the third of the situations outlined at the beginning of this section--whether the relationship is the same for different types of individuals. Perhaps the relationship between sex and voter preference varies with other characteristics of the individuals. Maybe among whites, females are more likely to prefer women candidates than the males are, but among African-Americans, there is little difference between males and females in terms of voter preference. This is the outcome shown in Table 3.4. This process is called **specification** because it specifies the conditions under which the relationship between sex and voter preference varies.

In the earlier section on bivariate analysis, we discussed the use of chi square. Remember that chi square is a test of independence used to determine if there is a relationship between two variables. Chi square is used in multivariate analysis the same way it is in bivariate analysis. There will be a separate value of chi square for each partial table in the

multivariate analysis. You should keep a number of warnings in mind. Chi square assumes that the expected frequencies for each cell are five or larger. As long as 80% of these expected frequencies are five or larger and no single expected frequency is very small, we don't have to worry. However, the expected frequencies often drop below five when the number of cases in a column or row gets too small. If this should occur, you will have to either **recode** (i.e., combine columns or rows) or eliminate a column or row from the table.

Table 3.4 -- Voting Preference by Sex Controlling for Race

	White			**African American**		
	Male %	Female %	Total %	Male %	Female %	Total %
Voting Preference						
Willing to Vote for a Woman	42.9	56.5	51.2	50.0	50.0	50.0
Not Willing to vote for a Woman	57.1	43.5	48.8	50.0	50.0	50.0
	100.00	100.00	100.00	100.00	100.00	100.00
	(310)	(490)	(800)	(50)	(50)	(100)

Another point to keep in mind is that chi square is affected by the number of cases in the table. With a lot of cases it is easy to reject the null hypothesis of no relationship. With a few cases, it can be quite hard to reject the null hypothesis. Also, consider the percentages within the table. Look for patterns. Do not rely on any single piece of information. Look at the whole picture.

We have concentrated on crosstabulation and chi square. There are other types of statistical analysis such as regression. When you have mastered these techniques, look at some other types of analysis.

Methods of Social Research

- Earl Babbie, 2007, *The Practice of Social Research* (11th edition), Thomson Wadsworth.

- Kenneth R. Hoover and Todd Donovan, 2004, *The Elements of Social Scientific Thinking* (8th Ed.). Thomson Wadsworth.

Survey Research and Sampling

- Earl Babbie, 2007, *The Practice of Social Research* (11th edition), Thomson Wadsworth.
- Arlene Fink, Linda Bourque, Eve P. Fieldler, Sabrina Mertens Oishi, Mark S. Litwin, 2003, The Survey Kit, Sage.
- Don A. Dillman, 2000, Mail and Internet Surveys – The Tailored Design Method (2nd edition), John Wiley & Sons.

Statistical Analysis

- Herman J. Loether and Donald G. McTavish, 1993, *Descriptive and Inferential Statistics An Introduction* (4th edition), Allyn and Bacon.
- J. Richard Kendrick, Jr., 2000, *Social Statistics*, Mayfield.
- Marija J. Norusis, 2005, *SPSS 13.0 Guide to Data Analysis*, Prentice Hall.

Elaboration and Causal Analysis

- Hirschi, Travis and Hanan C. Selvin, 1967, *Delinquency Research--An Appraisal of Analytic Methods*, Free Press.
- Rosenberg, Morris, 1968, *The Logic of Survey Analysis*, Basic Books.

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Chapter 4
Exercises Using the General Social Survey to Explore Relationships Among Variables

Note to the instructor: The data set used in this exercise is gss_04.por which consists of a subset of the 2004 General Social Survey. (Some of the variables in the GSS have been recoded to make them easier to use and some new variables have been created.) This exercise uses RECODE and CROSSTABS in SPSS to explore the relationships among variables. In CROSSTABS, students are asked to use percentages, chi square, and an appropriate measure of association. A good reference on using SPSS is *SPSS for Windows Version 13: A Basic Tutorial* by Linda Fiddler, Laura Hecht, Edward Nelson, Elizabeth Nelson, and Jim Ross. To order this book, call McGraw-Hill at 1-800-338-3987. The ISBN is 0-07-353671-7. You have permission to use this exercise and to revise it to fit your needs. Please send a copy of any revision to the authors.

Authors:
Ed Nelson and Elizabeth Nelson
Department of Sociology
California State University, Fresno
Fresno, CA 93740

Phone: 209-278-2275 (Ed)

Email: ednelson@csufresno.edu and/or elizn@csufresno.edu

Please contact the authors for additional information.

There are many social issues that you can explore using the General Social Survey. In these exercises, we are going to focus on abortion, confidence in institutions, tolerance for people holding unpopular ideas, and the types of social problems that people are willing to spend money on.

1. Seven variables focus on people's feelings about abortion: ABANY, ABDEFECT, ABHLTH, ABNOMORE, ABPOOR, ABRAPE, ABSINGLE. Each question asks respondents if

they think a woman ought to be able to obtain a legal abortion under varying circumstances. Choose one of these variables and use it as your dependent variable. Now choose one of the following variables as your independent variable: gender (SEX), class (CLASS), political party (PARTYID), and religion (RELIG or RELITEN or ATTEND or PRAY). If you want to use age, education or family income, then use ones of the variables that have already been recoded (AGE1, AGE2, EDUC1, INCOME981). Get the crosstabulation of these two variables along with the appropriate percentages and chi square and an appropriate measure of association (Gamma or Cramer's V). Write a short paragraph interpreting the relationship using the percentages and the other statistics to help you.

2. Several variables measure the amount of confidence the respondent has in the major institutions of our society. These include the military, big business, organized religion, education, the Executive Branch of the Federal Government, Congress, the press, and others. These variables all start with CON and there are thirteen of them.
    a. Run a frequency distribution to see which institutions respondents have the most confidence in and which they have the least confidence in. Be sure to use the appropriate percentages to make these comparisons. (Do you want the percents or the valid percents? Why?) Write a brief paragraph summarizing your results.
    b. Some people have more confidence in these institutions than others. Let's use political party preference (PARTYID) to divide respondents into Democrats, independents, and Republicans. You will have to recode PARTYID into three groups to do this. Combine strong and not strong Democrats into one group, combine strong and not strong Republicans into a second group and combine independents (near Democrat, near Republican, and independents) into a third group. Since there aren't many in the other category, let's recode "other" as a missing value so it will be removed from the table. If you recode it as a system-missing value, it will automatically be defined as a missing value.
    c. Now choose one of the social institutions that you think Democrats, Republicans and independents will have different levels of confidence in. Decide which is the independent and dependent variable

and get the crosstabulation. Be sure to ask for the appropriate percents, chi square, and measure of association. Write a paragraph indicating which group has the most confidence in this institution and which group has the least confidence. Use the percents, chi square, and measure of association to help you explain the relationship of these two variables.

3. Three sets of questions ask respondents whether they are tolerant of people who hold unpopular viewpoints. One set of questions asks respondents if they would allow five different types of people to teach in a college or university (COLATH, COLCOM, COLHOMO, COLMIL, COLRAC). Another set asks respondents if a book written by these five different types of people should be allowed in the public library (LIBATH, LIBCOM, LIBHOMO, LIBMIL, LIBRAC). Still another set asks respondents if they should be allowed to make a public speech in their community (SPKATH, SPKCOM, SPKHOMO, SPKMIL, SPKRAC). The five groups of people are those who are against churches and religion, communists, homosexuals, people who advocate doing away with elections and letting the military run the country, and those who claim Blacks are inferior.

   These variables have been combined into five other variables that measure tolerance for atheists, communists, homosexuals, militarists, and racists. Each variable is the sum of the three variables from the larger set of variables. For example, tolerance for racists is the sum of COLRAC, LIBRAC, and SPKRAC. Since each variable is coded 1 and 2, where 1 is the tolerant response and 2 is the intolerant response, the new variable (called TOLRAC) will vary from 3 to 6. The value 3 means that the respondent would be tolerant of racists in all three scenarios, while the value 6 means that the respondent would not be tolerant of racists in any of the three scenarios. The values 4 and 5 would be intermediate values.

   Get the frequency distributions for TOLATH, TOLCOM, TOLHOMO, TOLMIL, and TOLRAC to see if there is more tolerance for some of the groups than for others. Write a short paragraph explaining the results using the appropriate percents. (Be careful to decide whether you want the percents or the valid percents.)

Which groups of people would you expect to be more tolerant of homosexuals: men or women, Democrats or Republicans or independents, those living in the South or the Northeast or the Midwest or the West, working/lower class or middle/upper class? Choose one of these groupings and write a hypothesis that indicates your expectations. Write a short paragraph indicating why you think one group will be more tolerant of homosexuals than another.

Now find the variable in the list of variables that you want to use as the independent variable to test your hypothesis. Your dependent variable will be TOLHOMO. Get the crosstabulation to test your hypothesis. Be sure to get the appropriate percents, chi square, and measure of association. Write a short paragraph using the results to indicate whether the data support your hypothesis.

Chose one other tolerance variable (TOLATH, TOLCOM, TOLMIL, and TOLRAC) and repeat the analysis described above.

4. Americans decide what types of social problems to spend money on. The General Social Survey includes a series of questions that ask respondents whether we are spending too much, too little, or about the right amount of money on a series of problems. These problems include foreign aid, the military, big cities, crime, drugs, education, the environment, welfare, health, mass transportation, parks and recreation, the conditions of African-Americans, highways and bridges, social security, and space exploration.

   The General Social Survey includes two versions of most of these questions. All the spending variables start with NAT. The alternative version of each question ends with Y. For example, the questions on welfare are NATFARE and NATFAREY. NATFARE asks whether respondents think we are spending too much, too little, or about the right amount of money on "welfare." NATFAREY substitutes "assistance to the poor" for "welfare" in the question. A few questions have only one version of the question (i.e., no version Y). For this exercise, we will be using the original version of each question (i.e., the one that does not end in Y).

   Using the data, find out which problems respondents are the most likely to think we are spending too much money

on and which problems respondents think we are spending too little on. Write a brief paragraph summarizing your findings.

Republicans, Democrats, and independents often differ in terms of the problems they think we should be spending money on. Crosstabulate political party preference (PARTYID) and the spending variables to find out which problems Democrats think we should be spending more on, which problems Republicans think are more important, and which problems independents want to spend more money on. You will need to recode PARTYID into a smaller number of categories. See question 2 above for one way of recoding PARTYID. Use the appropriate percents, chi square, and measure of association to help you in your analysis. Write a short paragraph describing your results.

Class is another variable that often divides people on spending priorities. Use the variable CLASS to see if different classes have different spending priorities. You will have to recode CLASS. Do this by combining lower and working class into one category and middle and upper class into another category. Write a paragraph summarizing your results.

Find one variable where there are significant class and party differences on spending priorities. Which is more important--class or party? Think about how you are going to decide this. You will have to run a three-variable table to see the effect of one of these variables holding the other constant. For example, crosstab one of the NAT variables with class holding party constant. Then crosstab the same NAT variable with party holding class constant. Were there larger differences for class or for party when the other variable was held constant? Or were the differences about the same? Be sure to use the appropriate percents, chi square, and measure of association to help you. Write a paragraph or two describing your findings.

5. We already described the two different versions of the NAT variables. Why do you think the researchers did this? Different forms of the same questions often produce different results. For example, studies have found that people are more likely to say they would not allow something than they are to say they would forbid the same activity. The NAT... and the NAT...Y variables allow

us to study the effect of question wording on what respondents tell us.

Choose several pairs of NAT variables (e.g., NATFARE and NATFAREY, NATCRIME and NATCRIMY). Do a frequency distribution for both variables in the pair and see if the wording of the question makes any difference in the way respondents answered the questions. Keep looking at pairs of NAT variables until you find one pair where the question wording did make a difference and one pair where it didn't make a difference. Write a short paragraph summarizing your results.

It's possible that question wording makes more of a difference for some respondents than for others. Choose the pair of NAT variables where question wording did make a difference. Check to see if the wording of the questions made more of a difference for respondents with less education than for those with more education. How would you do this? You could choose DEGREE (highest degree) or EDUC (years of school completed) as your measure of education. If you choose EDUC, recode it into three or four categories. Then crosstab the NAT variables with your measure of education. Now compare the way respondents answered the NAT questions for each level of education. Try to construct a graph showing the differences. One way to do this would be to construct two line graphs. Each line graph could show the percent who felt we should be spending less money along the vertical axis and level of education along the horizontal axis. You would need two graphs--one for each of the NAT questions. These two line graphs could be placed on the same graph. Write a brief paragraph summarizing your results.

6. Other variables in the data set focus on women's issues and on issues of race. Most of the variables that begin with FE (FECHLD, FEFAM, FEPOL, FEPRESCH) deal with women's issues and variables that start with RAC (RACLIVE, RACOPEN) focus on race. Decide which of these issues you want to study and then look carefully at the appropriate variables for that issue in the codebook. Choose one variable that you would like to study. This will be your dependent variable. Now choose two independent variables that you think will be related to your dependent variable. For each variable, write a hypothesis that clearly states the relationship you expect to find between your independent and dependent variable. Indicate why you think this

hypothesis will be true. Get the crosstabulation that you need to test this hypothesis. Be sure to ask for the appropriate percents, chi square, and measure of association. Write a short paper that includes the hypothesis, the rationale for the hypothesis, the crosstabulation to test the hypothesis, and your interpretation of the table. Be sure to indicate whether the data support the hypothesis.

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Chapter 5
Research Design and Methods of Analysis for Change Over Time

Many of the problems that social scientists study are dynamic in character. A sociologist interested in the process of friendship formation would want to look at the way friendship dyads are formed over time. A political scientist interested in voting behavior would look at the ways in which people make up their minds for whom to vote. An economist studying consumer optimism might want to trace consumer sentiment over a period of years and see how that relates to other changes in the economy and the society.

A survey that looks at a cross-section of a group at one point in time (often called a cross-sectional survey) is not well suited to a study of dynamic processes. The political scientist interested in voting behavior might include questions that asked respondents who they had intended to vote for at the beginning of the campaign and who they plan to vote for now. This would allow an analysis of how voting decisions have changed over the campaign, but at a severe cost. The respondents' memory of who they preferred at the beginning of the campaign might be influenced by who they prefer later. We need some way to adapt cross-sectional surveys to the study of dynamic problems in which the focus is on change over time.

Cross-sectional studies focus on a group at one point in time. The decennial Census is a good example of a cross-sectional study. The 2000 Census describes the U.S. population at one point in time -- April, 2000. **Longitudinal studies** focus on the group at two or more points in time. We're going to look at three types of longitudinal studies -- the trend study, the panel study, and the cohort study.

TREND STUDIES

Trend studies use cross-sections at two or more points in time to examine change over time. The Virginia Slims Opinion Poll asked various questions about women at six points in time. The first poll was conducted in 1970, and it was repeated in 1972, 1974, 1980, 1985, and 1990. The Virginia Slims polls are

national probability samples of all adults living in the United States. These six cross-sectional surveys can be compared to trace changes in opinions and attitudes about women from 1970 to 1990.

Figure 5.1 shows the percent of women who favor efforts to strengthen and change women's status in society. The percent who favor such efforts increased steadily from 1970, while the percent who oppose decreased.

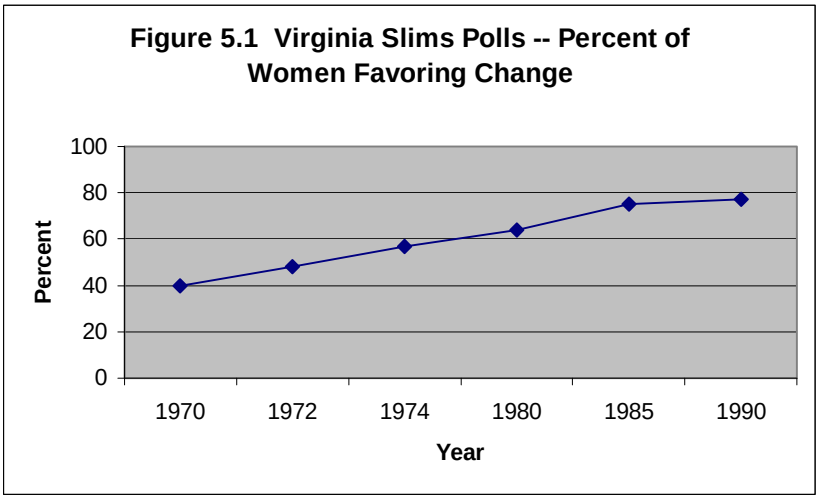
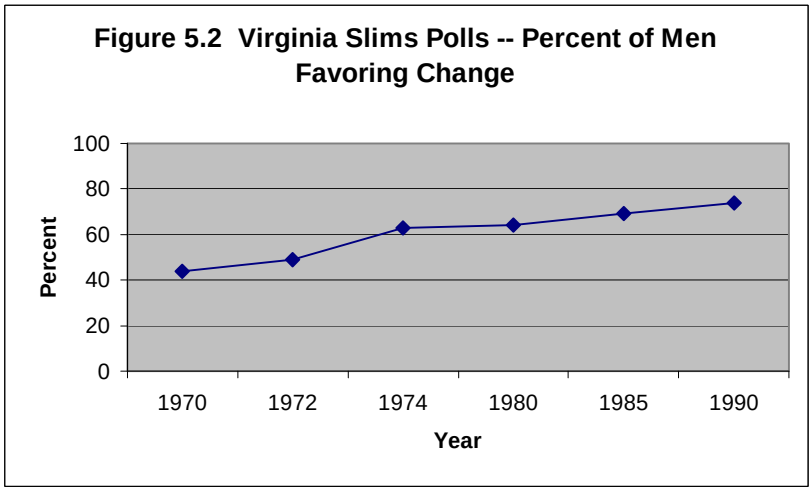
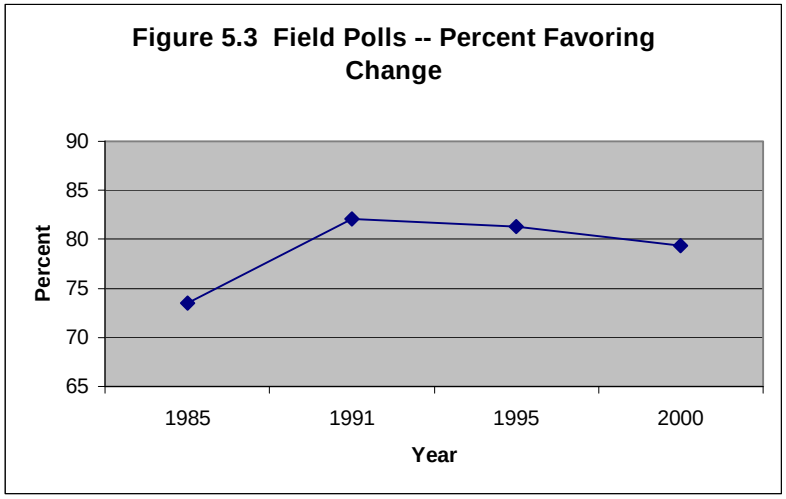


Figure 5.2 shows the same information for men. The pattern is the same for men as it is for women, but the percent of women favoring change in 1990 is slightly larger (77%) than the percent of men (74%).

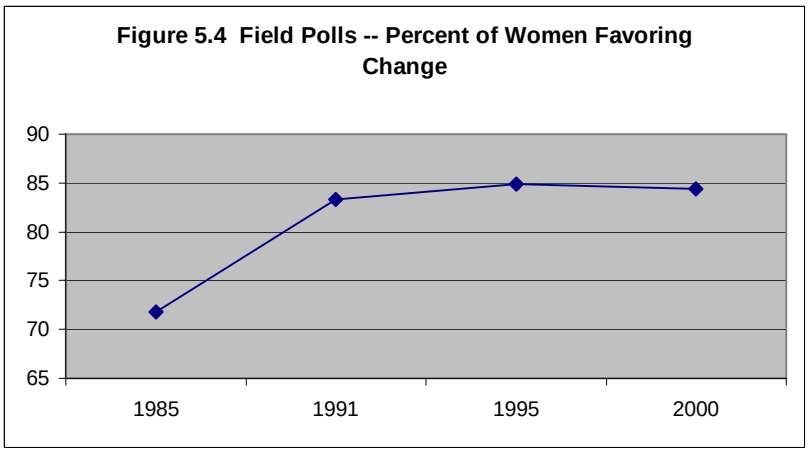


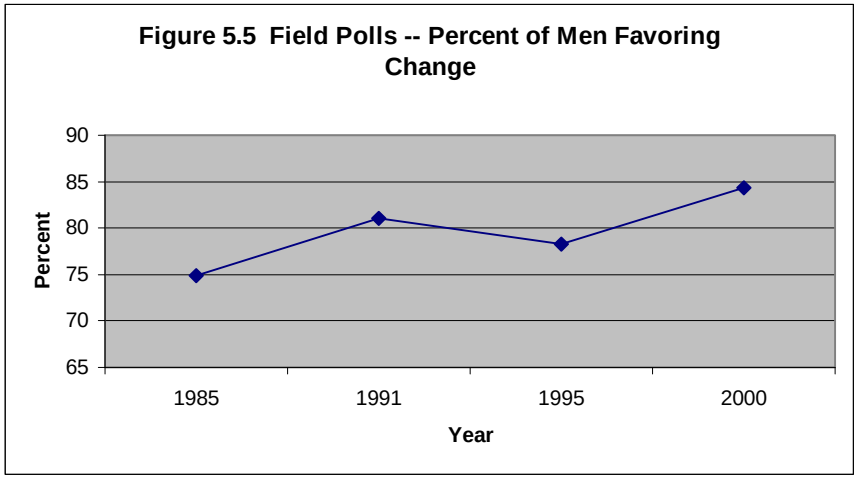
In 1985, 1991, 1995, and 2000, the Field Poll asked virtually the same question of a sample of California residents age 18

and over. Figure 5.3 shows the percent of respondents who favor such efforts. In California, there was an increase from 1985 to 1991, but there was very little change between 1991 and 2000.

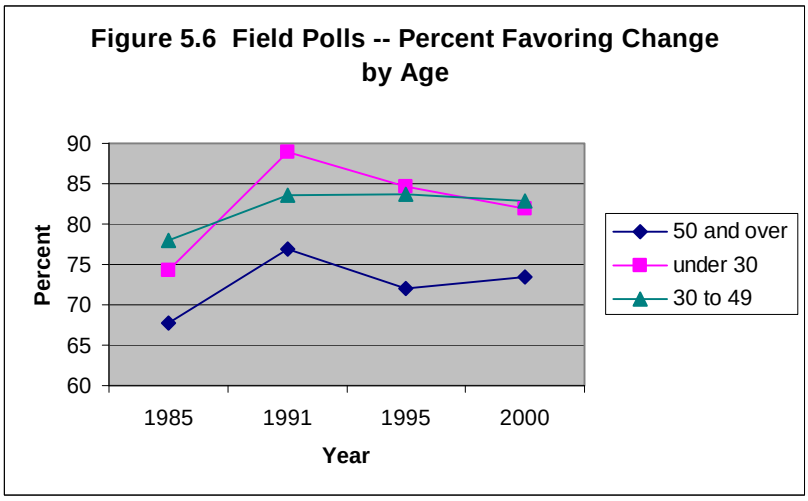


The Virginia Slims Polls showed the percentages separately for women and men. We can get this information for the Field Polls by crosstabulating the dependent variable by time and holding sex constant. This will give us two sets of figures, one for women and one for men. Figure 5.4 shows the percent of women who favor changes in women's roles, while Figure 5.5 shows the same for men. Notice that the percentages increased for women from 71.8% in 1985 to 83.3% in 1991 and then stayed fairly constant at 84.9% in 1995 and 84.4% in 2000, while the percentages for men increased from 74.9% in 1985 to 81.1% in 1991 and then decreased to 78.3% in 1995, but increased to 84.3% in 2000.





Let's look at the Field Polls for another example of change from 1985 to 2000. Figure 5.6 shows the percent who favor strengthening the status of women for three age groups--those under 30, those 30 to 49, and those 50 and over. We can learn several things from this table. First, the percent who favor such change increased in all age groups from 1985 to 1991 and then stayed about the same in the other three time periods. Second, in all time periods, those 50 and over are least likely to favor such change.



You might have noticed one difference between the Virginia Slims example and the Field Poll example. The Virginia Slims data are based on six time periods, while the Field Poll data are based on four time periods. Does the number of time periods make any difference? Figure 5.3 shows the percent who favor efforts to strengthen the status of women in 1985 and 1991 in the Field Poll. Clearly the percent who favor increased from 1985 to 1991. What if these were the only time periods for

which we have data? Let's imagine what the pattern might look like if we had data from four time periods prior to 1985. Figure 5.7 shows a long-term trend of increasing support. Figure 5.8 shows a long-term trend of little change with an increase from 1985 to 1991. Figure 5.9 shows a long-term trend of decreasing support with a reversal in this trend from 1985 to 1991. Without more time periods it is very difficult to determine what the long-term trend actually looks like. The 1995 and 2000 Field Polls indicate that there has been little change from 1991 to 2000. This example suggests that many time periods are better than few time periods. However, remember that every trend analysis must start with two points in time and build from that to a longer series.

Figure 5.7 Hypothetical Trend for Field Polls
Percent Favoring Change

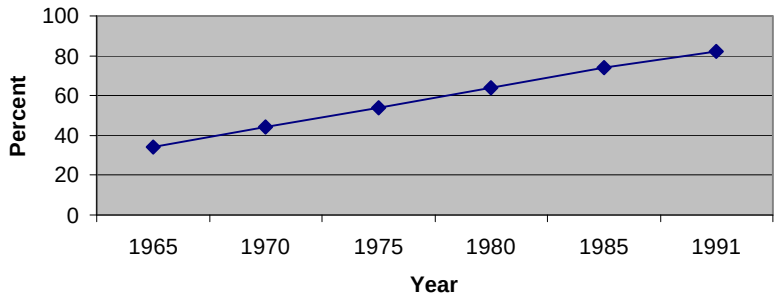
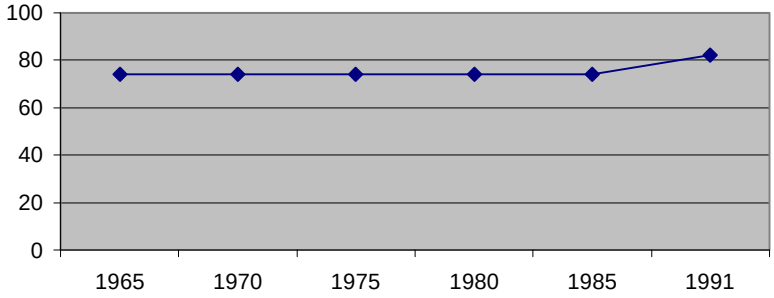


Figure 5.8 Hypothetical Trend for Field Polls
Percent Favoring Change



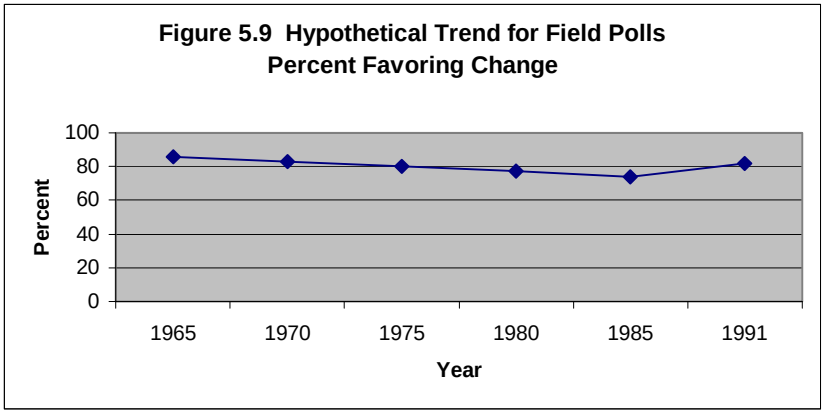


Figure 5.3 indicates that there was a 8.6 percentage point increase from 1985 to 1991 in the percent of respondents who favor strengthening the status of women. Was this produced by the shift of some individuals who opposed such changes in 1985 (or had no opinion) to a position that favors these changes in 1991? While it might seem tempting to accept this interpretation, it is not necessarily true. Table 5.1 shows a hypothetical example that is consistent with this interpretation. All of the individuals who favored change in 1985 also favor change in 1991. However, some of those who opposed change in 1985 now favor change in 1991 (n = 26), while more have shifted to a position of don't know (n = 79). Table 5.2 shows an example in which there has been a considerable shift in opinion in both directions--from oppose to favor and from favor to oppose. A considerable number of those who favored change in 1985 oppose change in 1991 (n = 100), while many of those who opposed change in 1985 now favor change (n = 115) or say they don't know (n = 90). Unfortunately, in a trend study we cannot choose between these alternatives. The most we can do in a trend study is to describe net changes between time periods. If we want to describe shifts from favor to oppose or from oppose to favor, then we need a different type of data called panel data.

PANEL STUDIES

Panel studies describe information about the same cases at two or more points in time. In a trend study we compare sample surveys describing the same population at two or more points in time. These samples consist of different cases. In a panel study we compare the same cases over time. If the Field Poll had been able to reinterview the same individuals in 1991 that were interviewed in 1985, then it would have been a panel study. The

advantage of panel studies is that we can choose between alternatives such as those presented in Tables 5.1 and 5.2. Panel data allow us to go beyond describing net changes between time periods. Panel data allow us to describe the types of shifts (e.g., from favor to oppose or from oppose to favor) that occur between time periods. We can also begin looking for factors that explain why some people change in one direction, while other individuals change in another direction, and still others do not change at all.

Table 5.1 -- Hypothetical Example Showing Opinion Shifts from 1985 to 1991

		1985			
		Favor	Oppose	Don't Know	Total
	Favor	684	26	25	735
1991	Oppose	0	141	0	141
	Don't Know	0	79	58	137
	Total	684	246	83	1013

Table 5.2 -- Hypothetical Example Showing Opinion Shifts from 1985 to 1991

		1985			
		Favor	Oppose	Don't Know	Total
	Favor	584	115	36	735
1991	Oppose	100	41	0	141
	Don't Know	0	90	47	137
	Total	684	246	83	1013

However, there are also problems with panel data. It is rare that all the cases are available in later time periods. This is called **panel mortality**. When the case is the individual, this may be because some individuals are not alive at a later point in time. However, all panel mortality may not be due to the death of respondents. Some respondents who cooperated initially may choose not to cooperate later. If particular types of individuals choose not to cooperate at a later point in time, then bias is introduced. For example, if low income respondents choose not to cooperate at a later point in time and if low income respondents are less likely to favor efforts to change the status of women, then part of the shift in opinion might be due to panel mortality.

Another problem with panel data is **reactivity**. If we ask people questions about the status of women at two or more points in time, the questioning process itself might produce opinion shifts. Perhaps the act of asking people about the status of women makes them more sensitive to women's issues. This increased sensitivity might mean they are more likely to favor or oppose changes in the status of women during later surveys. We call this reactivity because the respondents are reacting to the initial questioning.

COHORT STUDIES

Table 5.3 shows the percent who disagree that "women should take care of running their homes and leave running the country to men" by age for three of the General Social Surveys (1975, 1983, 1991). The General Social Survey is a national probability sample of all adults living in the United States. This table can be analyzed in several ways.

Table 5.3. Percent of Respondents Who Disagreed that "Women should take care of their homes and leave running the country up to men." *

Age	1975 (n)	1983 (n)	1991 (n)
18-25	78.5 (251)	87.2 (203)	85.0 (120)

26-33	75.6 (258)	84.7 (360)	90.1 (172)
34-41	67.5 (194)	83.0 (241)	91.5 (201)
42-49	68.3 (183)	81.4 (177)	86.5 (133)
50-57	58.4 (173)	76.3 (160)	78.0 (82)
58-65	56.1 (155)	66.5 (167)	80.2 (101)
66-73	39.0 (136)	57.6 (125)	58.0 (81)
74+	42.9 (91)	47.5 (99)	48.9 (90)
Total	64.4 (1441)	76.9 (1532)	80.8 (980)

*** The data are from the General Social Survey. The values inside the parentheses are the number of cases on which the percentages are based (i.e., the bases).**

First, we can compare opinions in the three time periods. We could compare the percent who disagree for each time period. This would show that there has been growing disagreement with this statement (64.4% of the total sample disagreed in 1975, 76.9% in 1983, and 80.8% in 1991). We could also compare the percent who disagree within each age group. For example, we could compare those age 34 to 41 in each of the three time periods (67.5% of those 34 to 41 disagree in 1975, 83.0% in 1983, and 91.5% in 1991) and repeat this for each of the age categories. This would involve comparing percentages across each row and this would show that there has been increasing disagreement even when we hold age constant.

Second, we could compare age categories within each time period. This would involve comparing percentages down within columns. This would show that the older respondents are less likely to disagree in each time period. In each of the three time periods, those age 66 to 73 and those 74 and over are considerably less likely to disagree than the younger respondents. (Note: even they increase over time--42.9% in 1975, 47.5% in 1983, 48.9% in 1991 for those 74 and over.)

Third, we could compare birth cohorts. Groups of people born within the same time period are called **birth cohorts**. The time period can be defined in any way that is appropriate for your analysis. Here we are using eight-year periods--all the people born within an eight-year period belong to the same birth cohort. Those who are 18 to 25 in 1975 would be 26 to 33 in 1983 and 34 to 41 in 1991. We could look at each of the birth cohorts in this table separately. This would involve comparing percentages along the diagonals running from the upper left part of the table to the lower right part. For example, for the birth cohort who was 18 to 25 in 1975, the percentages would be 78.5 in 1975, 84.7 in 1983, and 91.5 in 1991. In general, the cohorts are more likely to disagree with the statement in each successive time period. We could also compare other birth cohorts. The pattern described above is particularly noticeable for the four younger cohorts (i.e., those 18 to 25, 26 to 33, 34 to 41, and 42 to 49 in 1975).

Cohort studies follow one or more cohorts over a period of time. Cohort studies are usually based on two or more cross-sectional studies. In the example above, we have used three cross-sectional surveys (1975, 1983, 1991) and arranged the data so that we can compare birth cohorts.

SUMMARY

The purpose of this chapter has been to start you thinking about questions related to change over time. The analysis of such questions require longitudinal data. We have described three types of longitudinal studies -- the trend study, the panel study, and the cohort study.

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Chapter 6
Exercises Using 1975, 1982, 1989, 1996, 2002, and 2004 Subsets of the General Social Survey to Analyze Change Over Time

Note to the instructor: The data set used in this exercise is gss_75_82_89_96_02_04.por which consists of subsets of the 2004, 2002, 1996, 1989, 1982, 1975 General Social Surveys. (Some of the variables in the GSS have been recoded to make them easier to use and some new variables have been created.) This exercise uses RECODE and CROSSTABS in SPSS to explore the relationships among variables. In CROSSTABS, students are asked to use percentages, chi square, and an appropriate measure of association. A good reference on using SPSS is SPSS for Windows Version 13 A Basic Tutorial by Linda Fiddler, Laura Hecht, Edward Nelson, Elizabeth Nelson, and Jim Ross. To order this book, call McGraw-Hill at 1-800-338-3987. The ISBN is 0-07-353671-7. You have permission to use this exercise and to revise it to fit your needs. Please send a copy of any revision to the authors.

Authors:
Ed Nelson and Elizabeth Nelson
Department of Sociology
California State University, Fresno
Fresno, CA 93740

Phone: 209-278-2275 (Ed)

Email: ednelson@csufresno.edu and/or elizn@csufresno.edu

Please contact the authors for additional information.

These exercises will compare six cross-sectional surveys to analyze change over time. The exercises should be used with the accompanying codebook which is called sisscb.doc. Not all variables are available at all points in time. The codebook indicates which variables are available in which years. There is a variable called YEAR which indicates which year the data are from. The surveys in 1975, 1982, 1989, and 1996 were selected because they were seven years apart and contained basically

the same variables. There was no GSS survey in 2003, but there were surveys in 2002 and 2004. We suggest merging the 2002 and 2004 surveys and treating the merged sample as if it was the 2003 survey for the purpose of these exercises.

1. Seven variables focus on people's feelings about abortion: ABANY, ABDEFECT, ABHLTH, ABNOMORE, ABPOOR, ABRAPE, ABSINGLE. Each question asks respondents if they think a woman ought to be able to obtain a legal abortion under varying circumstances. Choose one of these variables to analyze change over time. (ABANY is not available for 1975.)

   What percent favor and oppose abortion for your variable in 1975? 1982? 1989? 1996? 2003? We'll need to use crosstabulation to get these figures. You will have to crosstabulate your variable by YEAR to do this. Be sure to ask for the column percents and chi square. What does chi square tell you about the difference between these percentages? (Note: If you combine the 2002 and 2004 surveys and treat the combined years as the equivalent to a 2003 survey, you will have to recode YEAR to combine 2002 and 2004 and then label it as 2003.)

   We want to discover which types of people are more likely to change. Let's start by asking if men or women are more likely to change. You will have to crosstabulate the abortion variable by YEAR by sex, asking for the column percents and chi square. Write a short paragraph describing changes for men and for women. Be sure to cite the appropriate percent differences and use chi square in your analysis.

   Now, find out if younger, middle-aged, or older respondents are more likely to change and if those with less education are more likely to change than those with more education. You'll need to recode age and education before running the crosstabs. Another possibility is to use the recodes we included in the data set (AGE1, AGE2, ECUC1). You could also use DEGREE which does not have to recoded. Write a short paragraph describing the results.

2. Several variables measure the amount of confidence the respondent has in the major institutions of our society. These include the military, big business, organized religion, education, the Executive Branch of the Federal Government, Congress, the press, and others. These

variables all start with CON and there are thirteen of them.

Choose one of these variables that you want to analyze over time. Crosstabulate your variable by YEAR and describe the changes that have occurred over time. Be sure to use the column percents and chi square in your analysis.

Some people have more confidence in these institutions than others. Let's use political party preference (PARTYID) to divide respondents into Democrats, independents, and Republicans. You will have to recode PARTYID into three groups to do this. Combine strong and not strong Democrats into one group, combine strong and not strong Republicans into a second group and combine independents (near Democrat, near Republican, and independents) into a third group. When you recode, use recoding into different variables and call your recoded variable PARTYID1.

Now let's analyze change over time for Democrats, independents, and Republicans separately. Crosstabulate your variable by YEAR by political party preference and describe the changes separately for each party. Use the column percents and chi square to help you in your analysis. Have there been greater changes for Democrats? for Republicans? for independents? Have the changes been in the same direction for all three groups?

3. Three sets of questions ask respondents whether they are tolerant of people who hold unpopular viewpoints. One set of questions asks respondents if they would allow five different types of people to teach in a college or university (COLATH, COLCOM, COLHOMO, COLMIL, COLRAC). Another set asks respondents if a book written by these five different types of people should be allowed in the public library (LIBATH, LIBCOM, LIBHOMO, LIBMIL, LIBRAC). Still another set asks respondents if they should be allowed to make a public speech in their community (SPKATH, SPKCOM, SPKHOMO, SPKMIL, SPKRAC). The five groups of people are those who are against churches and religion, communists, homosexuals, people who advocate doing away with elections and letting the military run the country, and those who claim African-Americans are inferior.

These variables have been combined into five other variables that measure tolerance for atheists, communists, homosexuals, militarists, and racists. Each variable is the sum of the three variables from the larger set of variables. For example, tolerance for racists is the sum of COLRAC, LIBRAC, and SPKRAC. Since each variable is coded 1 and 2, where 1 is the tolerant response and 2 is the intolerant response, the new variable (called TOLRAC) will vary from 3 to 6. The value 3 means that the respondent would be tolerant of racists in all three scenarios, while the value 6 means that the respondent would not be tolerant of racists in any of the three scenarios. The values 4 and 5 would be intermediate values.

Crosstabulate TOLATH, TOLCOM, TOLHOM, TOLMIL, and TOLRAC by YEAR to see the changes that have occurred over time for tolerance. (These variables are not available for 1975.) Use the column percents and chi square in your analysis. Write a brief paragraph describing your results.

Who would you expect to be more tolerant-men or women? To find out, crosstabulate the tolerance variables by sex by YEAR. This will give you the percents you need to compare men and women for each of the years. Did you get the same results for each year? Were the differences between the figures for men and women about the same for each year or was there considerable variation from year to year?

4. Americans decide what types of social problems to spend money on. The General Social Survey includes a series of questions that ask respondents whether we are spending too much, too little, or about the right amount of money on a series of problems. These problems include foreign aid, the military, big cities, crime, drugs, education, the environment, welfare, health, mass transportation, parks and recreation, the conditions of African-Americans, highways and bridges, social security, and space exploration.

   The General Social Survey includes two versions of most of these questions. All the spending variables start with NAT. The alternative version of each question ends with Y. For example, the questions on welfare are NATFARE and NATFAREY. NATFARE asks whether respondents think we are spending too much, too little, or about the right

amount of money on "welfare." NATFAREY substitutes "assistance to the poor" for "welfare" in the question. A few questions have only one version of the question (i.e., no version Y). For this exercise, we will be using the original version of each question (i.e., the one that does not end in Y).

Select one of the NAT variables and analyze the changes in opinion over time. Notice that a few of the NAT variables were not available in 1975 and 1982. Choose one of the NAT variables that was available in all time periods. Look at both the overall changes for the entire sample and the changes for subgroups (e.g., men and women, younger and older). Use chi square in your analysis. Write a brief report explaining what you did and describing the results. Include your tables in the report.

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Appendix A:
Codebook for the Social Issues Subset of the General Social Survey

This codebook accompanies gss_75_82_89_96_02_04.por and gss_04.por, which are subsets of the General Social Survey. gss_75_82_89_96_02_04.por includes data from all six time periods of the GSS—2004, 2002, 1996, 1989, 1982, 1975. gss_04 includes data from only the 2004 GSS. Some of the variables in the GSS have been recoded to make them easier to use and some new variables have been created. The variable "Year" indicates the year of the survey.

List of Variables

Variable Name	Years	Variable Label
ABANY	04 02 96 89 82	ABORTION--FOR ANY REASON
ABDEFECT	04 02 96 89 82 75	ABORTION--STRONG CHANCE SERIOUS DEFECT
ABHLTH	04 02 96 89 82 75	ABORTION--WOMANS HEALTH ENDANGERED
ABNOMORE	04 02 96 89 82 75	ABORTION--MARRIED,NO MORE CHILDREN
ABPOOR	04 02 96 89 82 75	ABORTION--LOW INCOME, CANT AFFORD MORE
ABRAPE	04 02 96 89 82 75	ABORTION--PREGNANT AS RESULT OF RAPE
ABSINGLE	04 02 96 89 82 75	ABORTION--NOT MARRIED
ADULTS	04 02 96 89 82 75	NUMBER OF ADULTS IN HOUSEHOLD
AGE	04 02 96 89 82 75	AGE OF RESPONDENT
AGE1	04 02 96 89 82 75	AGE RECODED IN TEN YEAR INTERVALS
AGE2	04 02 96 89 82 75	AGE RECODED IN THREE CATEGORIES
AGED	04 02 96 89 75	SHOULD AGED LIVE WITH THEIR CHILDREN
ATTEND	04 02 96 89 82 75	HOW OFTEN R ATTENDS RELIGIOUS SERVICES
CHILDS	04 02 96	NUMBER OF CHILDREN

	89 82 75	
CLASS	04 02 96 89 82 75	SUBJECTIVE CLASS IDENTIFICATION
COLATH	04 02 96 89 82	ALLOW ANTI-RELIGIONIST TO TEACH
COLCOM	04 02 96 89 82	ALLOW COMMUNIST TO TEACH
COLHOMO	04 02 96 89 82	ALLOW HOMOSEXUAL TO TEACH
COLMIL	04 02 96 89 82	ALLOW MILITARIST TO TEACH
COLRAC	04 02 96 89 82	ALLOW RACIST TO TEACH
CONARMY	04 02 96 89 82 75	CONFIDENCE IN MILITARY
CONBUS	04 02 96 89 82 75	CONFIDENCE IN MAJOR COMPANIES
CONCLERG	04 02 96 89 82 75	CONFIDENCE IN ORGANIZED RELIGION
CONEDUC	04 02 96 89 82 75	CONFIDENCE IN EDUCATION
CONFED	04 02 96 89 82 75	CONFIDENCE IN EXEC BRANCH OF FED GOVT
CONFINAN	04 02 96 89 82 75	CONFIDENCE IN BANKS & FINANCIAL INST
CONJUDGE	04 02 96 89 82 75	CONFIDENCE IN U.S. SUPREME COURT
CONLABOR	04 02 96 89 82 75	CONFIDENCE IN ORGANIZED LABOR
CONLEGIS	04 02 96 89 82 75	CONFIDENCE IN CONGRESS
CONMEDIC	04 02 96 89 82 75	CONFIDENCE IN MEDICINE
CONPRESS	04 02 96 89 82 75	CONFIDENCE IN PRESS
CONSCI	04 02 96 89 82 75	CONFIDENCE IN SCIENTIFIC COMMUNITY
CONTV	04 02 96 89 82 75	CONFIDENCE IN TELEVISION
COURTS	04 02 96 89 82 75	COURTS DEALING WITH CRIMINALS
DEGREE	04 02 96 89 82 75	R'S HIGHEST DEGREE
EDUC	04 02 96 89 82 75	HIGHEST YEAR OF SCHOOL COMPLETED
EDUC1	04 02 96 89 82 75	EDUCATION RECODED

EQWLTH	04 02 96 89	SHOULD GOVT REDUCE INCOME DIFFERENCES
FEAR	04 02 96 89 82	AFRAID TO WALK AT NIGHT IN NEIGHBORHOOD
FECHLD	04 02 96 89	MOTHER WORKING DOESNT HURT CHILDREN
FEFAM	04 02 96 89	BETTER FOR MAN TO WORK, WOMAN TEND HOME
FEPOL	04 02 96 89 82 75	WOMEN NOT SUITED FOR POLITICS
FEPRESCH	04 02 96 89	PRESCHOOL KIDS SUFFER IF MOTHER WORKS
FINALTER	04 02 96 89 82 75	CHANGE IN FINANCIAL SITUATION
FINRELA	04 02 96 89 82 75	OPINION OF FAMILY INCOME
FUND	04 02 96 89 82 75	HOW FUNDAMENTALIST IS R CURRENTLY
INCOME98	04 02	TOTAL FAMILY INCOME
INCOME91	96	TOTAL FAMILY INCOME
INCOME86	89	TOTAL FAMILY INCOME
INCOME82	82	TOTAL FAMILY INCOME
INCOME	75	TOTAL FAMILY INCOME
LIBATH	04 02 96 89 82	ALLOW ANTI-RELIGIOUS BOOK IN LIBRARY
LIBCOM	04 02 96 89 82	ALLOW COMMUNISTS BOOK IN LIBRARY
LIBHOMO	04 02 96 89 82	ALLOW HOMOSEXUALS BOOK IN LIBRARY
LIBMIL	04 02 96 89 82	ALLOW MILITARISTS BOOK IN LIBRARY
LIBRAC	04 02 96 89 82	ALLOW RACISTS BOOK IN LIBRARY
MARITAL	04 02 96 89 82 75	MARITAL STATUS
NATAID	04 02 96 89 82 75	FOREIGN AID
NATAIDY	04 02 96 89	ASSISTANCE TO OTHER COUNTRIES -- VERSION Y
NATARMS	04 02 96 89 82 75	MILITARY, ARMAMENTS, AND DEFENSE
NATARMSY	04 02 96 89	NATIONAL DEFENSE -- VERSION Y
NATCITY	04 02 96 89 82 75	SOLVING PROBLEMS OF BIG CITIES
NATCITYY	04 02 96	ASSISTANCE TO BIG CITIES -

	89	-VERSION Y
NATCRIME	04 02 96 89 82 75	HALTING RISING CRIME RATE
NATCRIMY	04 02 96 89	LAW ENFORCEMENT -- VERSION Y
NATDRUG	04 02 96 89 82 75	DEALING WITH DRUG ADDICTION
NATDRUGY	04 02 96 89	DRUG REHABILITATION -- VERSION Y
NATEDUC	04 02 96 89 82 75	IMPROVING NATIONS EDUCATION SYSTEM
NATEDUCY	04 02 96 89	EDUCATION -- VERSION Y
NATENVIR	04 02 96 89 82 75	IMPROVING&PROTECTING ENVIRONMENT
NATENVIY	04 02 96 89	THE ENVIRONMENT -- VERSION Y
NATFARE	04 02 96 89 82 75	WELFARE
NATFAREY	04 02 96 89	ASSISTANCE TO THE POOR -- VERSION Y
NATHEAL	04 02 96 89 82 75	IMPROVING&PROTECTING NATIONS HEALTH
NATHEALY	04 02 96 89	HEALTH -- VERSION Y
NATMASS	04 02 96 89	MASS TRANSPORTATION
NATPARK	04 02 96 89	PARKS AND RECREATION
NATRACE	04 02 96 89 82 75	IMPROVING THE CONDITIONS OF BLACKS
NATRACEY	04 02 96 89	ASSISTANCE TO BLACKS -- VERSION Y
NATROAD	04 02 96 89	HIGHWAYS AND BRIDGES
NATSOC	04 02 96 89	SOCIAL SECURITY
NATSPAC	04 02 96 89 82 75	SPACE EXPLORATION PROGRAM
NATSPACY	04 02 96 89	SPACE EXPLORATION -- VERSION Y
PARTYID	04 02 96 89 82 75	POLITICAL PARTY AFFILIATION
PRAY	04 02 96 89	HOW OFTEN DOES R PRAY
PRAYER	04 02 96 89 82 75	SUPPORT SUPREME COURT'S DECISION

PRESTG80	04 02 96 89	R'S OCCUPATIONAL PRESTIGE SCORE (1980)
PRESTIGE	82 75	R'S OCCUPATIONAL PRESTIGE SCORE (1970)
PRESTG1	04 02 96 89 82 75	RECODED OCCUPATIONAL PRESTIGE
RACE	96 89 82 75	RACE OF RESPONDENT
RACECEN1	04 02	RACE OF RESPONDENT
RACECEN2	04 02	RACE OF RESPONDENT
RACECEN3	04 02	RACE OF RESPONDENT
RACLIVE	04 02 96 89 82 75	ANY OPPOSITE RACE IN NEIGHBORHOOD
RACOPEN	04 96 89 75	VOTE ON OPEN HOUSING LAW
REGION	04 02 96 89 82 75	REGION OF INTERVIEW
RELIG	04 02 96 89 82 75	R'S RELIGIOUS PREFERENCE
RELGOSA	04 02 96 89	RELIGIOSITY (BASED ON ATTEND, PRAY, RELITEN)
RELGOSB	04 02 96 89 82 75	RELIGIOSITY (BASED ON ATTEND, RELITEN)
RELITEN	04 02 96 89 82 75	STRENGTH OF AFFILIATION
RINCOM98	04 02	RESPONDENT’S INCOME
RINCOM91	96	RESPONDENT'S INCOME
RINCOM86	89	RESPONDENT'S INCOME
RINCOM82	82	RESPONDENT'S INCOME
RINCOME	75	RESPONDENT'S INCOME
SATFIN	04 02 96 89 82 75	SATISFACTION WITH FINANCIAL SITUATION
SEI	04 02 96 89	RESPONDENT SOCIOECONOMIC INDEX
SEI1	04 02 96 89	RECODED SOCIOECONOMIC INDEX
SEX	04 02 96 89 82 75	RESPONDENT'S SEX
SPKATH	04 02 96 89 82	ALLOW ANTI-RELIGIONIST TO SPEAK
SPKCOM	04 02 96 89 82	ALLOW COMMUNIST TO SPEAK
SPKHOMO	04 02 96 89 82	ALLOW HOMOSEXUAL TO SPEAK
SPKMIL	04 02 96 89 82	ALLOW MILITARIST TO SPEAK

SPKRAC	04 02 96 89 82	ALLOW RACIST TO SPEAK
SUICIDE1	04 02 96 89 82	SUICIDE IF INCURABLE DISEASE
SUICIDE2	04 02 96 89 82	SUICIDE IF BANKRUPT
SUICIDE3	04 02 96 89 82	SUICIDE IF DISHONORED FAMILY
SUICIDE4	04 02 96 89 82	SUICIDE IF TIRED OF LIVING
TOLATH	04 02 96 89 82	TOLERANCE--ANTI RELIGIONIST
TOLCOM	04 02 96 89 82	TOLERANCE--COMMUNIST
TOLHOMO	04 02 96 89 82	TOLERANCE--HOMOSEXUAL
TOLMIL	04 02 96 89 82	TOLERANCE--MILITARIST
TOLRAC	04 02 96 89 82	TOLERANCE--RACIST
WEIGHT	04 02 96 89 82 75	WEIGHT VARIABLE
YEAR	04 02 96 89 82 75	YEAR OF SURVEY

ABANY Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion the woman wants it for any reason?

1	YES
2	NO
0	NAP
8	DK
9	NA

ABDEFECT Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if there is a strong chance of serious defect in the baby?

1	YES
2	NO
0	NAP
8	DK
9	NA

ABHLTH Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if the woman's own health is seriously endangered by the pregnancy?

1 YES
2 NO
0 NAP
8 DK
9 NA

ABNOMORE Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children?

1 YES
2 NO
0 NAP
8 DK
9 NA

ABPOOR Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if the family has a very low income and cannot afford any more children?

1 YES
2 NO
0 NAP
8 DK
9 NA

ABRAPE Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if she became pregnant as a result of rape?

1 YES
2 NO
0 NAP
8 DK
9 NA

ABSINGLE Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if she is not married and does not want to marry the man?

1 YES
2 NO

0	NAP
8	DK
9	NA

AGE Age of respondent in years

9	NA

AGE1 Age recoded ten year intervals

1	UNDER 30
2	30 TO 39
3	40 TO 49
4	50 TO 59
5	60 TO 69
6	70 AND OVER
9	MISSING

AGE2 Age recoded in three categories

1	UNDER 30
2	30 TO 49
3	50 AND OVER
9	MISSING

AGED As you know, many older people share a home with their grown children. Do you think this is generally a good idea or a bad idea?

1	A GOOD IDEA
2	DEPENDS
3	A BAD IDEA
0	NAP
8	DK
9	NA

ATTEND How often do you attend religious services?

0	NEVER
1	LT ONCE A YEAR
2	ONCE A YEAR
3	SEVERAL TIMES A YEAR
4	ONCE A MONTH

5 2-3X A MONTH
6 NEARLY EVERY WEEK
7 EVERY WEEK
8 MORE THAN ONCE WEEK
9 DK, NA

CHILDS How many children have you ever had? (Includes all that were born alive at any time and those from a previous marriage)

8 EIGHT OR MORE
9 NA

CLASS If you were asked to use one of four names for your social class, which would you say you belong in: the lower class, the working class, the middle class, or the upper class?

1 LOWER CLASS
2 WORKING CLASS
3 MIDDLE CLASS
4 UPPER CLASS
0 NAP
8 DK
9 NA

COLATH There are always some people whose ideas are considered bad or dangerous by other people. For instance, somebody who is against all churches and religion . . .
Should such a person be allowed to teach in a college or university, or not?

1 ALLOWED
2 NOT ALLOWED
0 NAP
8 DK
9 NA

COLCOM Now, I should like to ask you some questions about a man who admits he is a Communist. Suppose he is teaching in a college. Should he be allowed to teach, or not?

1 ALLOWED
2 NOT ALLOWED
0 NAP
8 DK

9	NA

COLHOMO And what about a man who admits that he is a homosexual? Should such a person be allowed to teach in a college or university, or not?

1	ALLOW
2	NOT ALLOW
0	NAP
8	DK
9	NA

COLMIL Consider a person who advocates doing away with elections and letting the military run the country. Should such a person be allowed to teach in a college or university, or not?

1	ALLOW
2	NOT ALLOW
0	NAP
8	DK
9	NA

COLRAC Or consider a person who believes that Blacks are genetically inferior. Should such a person be allowed to teach in a college or university, or not?

1	ALLOW
2	NOT ALLOW
0	NAP
8	DK
9	NA

CONARMY I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Military

1	A GREAT DEAL
2	ONLY SOME
3	HARDLY ANY
0	NAP
8	DK
9	NA

CONBUS I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Major companies

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONCLERG I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Organized religion

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONEDUC I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Education

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONFED I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Executive branch of the federal government

1 A GREAT DEAL

2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONFINAN I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Banks and financial institutions

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONJUDGE I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

U.S. Supreme Court

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONLABOR I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Organized labor

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONLEGIS I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Congress

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONMEDIC I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Medicine

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONPRESS I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Press

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONSCI I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

Scientific Community

1 A GREAT DEAL

2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

CONTV I am going to name some institutions in this country. As far as the people running these institutions are concerned, would you say you have a great deal of confidence, only some confidence, or hardly any confidence at all in them?

TV

1 A GREAT DEAL
2 ONLY SOME
3 HARDLY ANY
0 NAP
8 DK
9 NA

COURTS In general, do you think the courts in this area deal too harshly or not harshly enough with criminals?

1 TOO HARSH
2 ABOUT RIGHT
3 NOT HARSH ENOUGH
8 DK
9 NA

DEGREE Respondent's highest degree

0 LT HIGH SCHOOL
1 HIGH SCHOOL
2 JUNIOR COLLEGE
3 BACHELOR
4 GRADUATE
7 NAP
8 DK
9 NA

EDUC What is the highest grade in elementary school or high school that you finished and got credit for?

8 DK
9 NA

EDUC1 Education recoded

1 12 YEARS OR LESS
2 13 TO 15 YEARS
3 16 YEARS OR MORE
9 DK, NA

EQWLTH Some people think that the government in Washington ought to reduce the income differences between the rich and the poor, perhaps by raising the taxes of wealthy families or by giving income assistance to the poor. Others think that the government should not concern itself with reducing this income difference between the rich and the poor.

1 GOVERNMENT REDUCE DIFFERENCES
7 NO GOVERNMENT ACTION
0 NAP
8 DK
9 NA

FEAR Is there any area right around here--that is, within a mile--where you would be afraid to walk alone at night?

1 YES
2 NO
0 NAP
8 DK

FECHLD A working mother can establish just as warm and secure a relationship with her children as a mother who does not work.

1 STRONGLY AGREE
2 AGREE
3 DISAGREE
4 STRONGLY DISAGREE
0 NAP
8 DK
9 NA

FEFAM It is much better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family.

1 STRONGLY AGREE
2 AGREE
3 DISAGREE
4 STRONGLY DISAGREE

0 NAP
8 DK
9 NA

FEPOL Most men are better suited emotionally for politics than are most women.

1 AGREE
2 DISAGREE
0 NAP
8 NOT SURE
9 NA

FEPRESCH A preschool child is likely to suffer if his or her mother works.

1 STRONGLY AGREE
2 AGREE
3 DISAGREE
4 STRONGLY DISAGREE
0 NAP
8 DK
9 NA

FINALTER During the last few years, has your financial situation been getting better, worse, or has it stayed the same?

1 BETTER
2 STAYED SAME
3 WORSE
8 DK
9 NA

FINRELA Compared with American families in general, would you say your family income is far below average, below average, average, above average, or far above average?

1 FAR BELOW AVERAGE
2 BELOW AVERAGE
3 AVERAGE
4 ABOVE AVERAGE
5 FAR ABOVE AVERAGE
8 DK
9 NA

FUND Fundamentalism/Liberalism of respondent's religion

1 FUNDAMENTALIST
2 MODERATE
3 LIBERAL
9 NA

INCOME98 In which of these groups did your total family income, from all sources, fall last year before taxes? [see footnote 1]

1 Under $10000
2 $10000-14999
3 $15000-19999
4 $20000-24999
5 $25000-34999
6 $35000-49999
7 $50000-59999
8 $60000-74999
9 $75000-89999
10 $90000-109999
11 $110000+

INCOME91 In which of these groups did your total family income, from all sources, fall last year before taxes? [see footnote 1]

1 Under $10000
2 $10000-14999
3 $15000-19999
4 $20000-24999
5 $25000-34999
6 $35000-49999
7 $50000-59999
8 $60000-74999
9 $75000+

INCOME86 [see footnote 1]

1 Under $10000
2 $10000-14999
3 $15000-19999
4 $20000-24999
5 $25000-34999
6 $35000-49999
7 $50000-59999
8 $60000+

INCOME82 [see footnote 1]

1	Under $10000
2	$10000-14999
3	$15000-19999
4	$20000-24999
5	$25000-34999
6	$35000-49999
7	$50000+

INCOME [see footnote 1]

1	Under $10000
2	$10000-14999
3	$15000-19999
4	$20000-24999
5	$25000+

LIBATH There are always some people whose ideas are considered bad or dangerous by other people. For instance, somebody who is against all churches and religion . . . If some people in your community suggested that a book he wrote against churches and religion should be taken out of your public library, would you favor removing this book, or not?

1	NOT REMOVE
2	REMOVE
0	NAP
8	DK
9	NA

LIBCOM Now, I should like to ask you some questions about a man who admits he is a Communist. Suppose he wrote a book which is in your public library. Somebody in your community suggests that the book should be removed from the library. Would you favor removing it, or not?

1	NOT REMOVE
2	REMOVE
0	NAP
8	DK
9	NA

LIBHOMO And what about a man who admits that he is a homosexual? If some people in your community suggested that a book he wrote in favor of homosexuality should be taken out of your public library, would you favor removing this book, or not?

1 NOT REMOVE
2 REMOVE
0 NAP
8 DK
9 NA

LIBMIL Consider a person who advocates doing away with elections and letting the military run the country. Suppose he wrote a book advocating doing away with elections and letting the military run the country. Somebody in your community suggests that the book be removed from the public library. Would you favor removing it, or not?

1 NOT REMOVE
2 REMOVE
0 NAP
8 DK
9 NA

LIBRAC Or consider a person who believes that Blacks are genetically inferior. If some people in your community suggested that a book he wrote which said Blacks are inferior should be taken out of your public library, would you favor removing this book, or not?

1 NOT REMOVE
2 REMOVE
0 NAP
8 DK
9 NA

MARITAL Are you currently--married, widowed, divorced, separated, or have you never been married?

1 MARRIED
2 WIDOWED
3 DIVORCED
4 SEPARATED
5 NEVER MARRIED
9 NA

NATAID 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. Are we spending too much money, too little money, or about the right amount on...

Foreign aid

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATAIDY 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. Are we spending too much money, too little money, or about the right amount on...

Assistance to other countries

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATARMS 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. Are we spending too much money, too little money, or about the right amount on...

The military, armaments and defense

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATARMSY 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. Are we spending too much money, too little money, or about the right amount on...

National defense

1 TOO LITTLE

2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATCITY 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. Are we spending too much money, too little money, or about the right amount on...

Solving the problems of the big cities

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATCITYY 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. Are we spending too much money, too little money, or about the right amount on...

Assistance to big cities

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATCRIME 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. Are we spending too much money, too little money, or about the right amount on...

Halting the rising crime rate

1 TOO LITTLE
2 ABOUT RIGHT

3 TOO MUCH
0 NAP
8 DK
9 NA

NATCRIMY 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. Are we spending too much money, too little money, or about the right amount on...

Law enforcement

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATDRUG 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. Are we spending too much money, too little money, or about the right amount on...

Dealing with drug addiction

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATDRUGY 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. Are we spending too much money, too little money, or about the right amount on...

Drug rehabilitation

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH

0 NAP
8 DK
9 NA

NATEDUC 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. Are we spending too much money, too little money, or about the right amount on...

Improving the nation's education system

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATEDUCY 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. Are we spending too much money, too little money, or about the right amount on...

Education

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATENVIR 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. Are we spending too much money, too little money, or about the right amount on...

Improving and protecting the environment

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP

8 DK
9 NA

NATENVIY 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. Are we spending too much money, too little money, or about the right amount on...

The environment

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATFARE 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. Are we spending too much money, too little money, or about the right amount on...

Welfare

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATFAREY 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. Are we spending too much money, too little money, or about the right amount on...

Assistance to the poor

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK

9 NA

NATHEAL 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. Are we spending too much money, too little money, or about the right amount on...

Improving and protecting the nation's health

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATHEALY 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. Are we spending too much money, too little money, or about the right amount on...

Health

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATMASS 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. Are we spending too much money, too little money, or about the right amount on...

Mass Transportation

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATPARK 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. Are we spending too much money, too little money, or about the right amount on...

Parks and recreation

1	TOO LITTLE
2	ABOUT RIGHT
3	TOO MUCH
0	NAP
8	DK
9	NA

NATRACE 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. Are we spending too much money, too little money, or about the right amount on...

Improving the conditions of Blacks

1	TOO LITTLE
2	ABOUT RIGHT
3	TOO MUCH
0	NAP
8	DK
9	NA

NATRACEY 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. Are we spending too much money, too little money, or about the right amount on...

Assistance to blacks

1	TOO LITTLE
2	ABOUT RIGHT
3	TOO MUCH
0	NAP
8	DK
9	NA

NATROAD 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. Are we spending too much money, too little money, or about the right amount on...

Highways and bridges

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATSOC 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. Are we spending too much money, too little money, or about the right amount on...

Social Security

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATSPAC 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)?

Space exploration program

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

NATSPACY 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)?

Space exploration

1 TOO LITTLE
2 ABOUT RIGHT
3 TOO MUCH
0 NAP
8 DK
9 NA

PARTYID Generally speaking, do you usually think of yourself as a Republican, Democrat, Independent, or what?

0 STRONG DEMOCRAT
1 NOT STRONG DEMOCRAT
2 INDEPENDENT, NEAR DEMOCRAT
3 INDEPENDENT
4 INDEPENDENT, NEAR REPUBLICAN
5 NOT STRONG REPUBLICAN
6 STRONG REPUBLICAN
7 OTHER PARTY
9 NA

PRAY About how often do you pray?

1 SEVERAL TIMES A DAY
2 ONCE A DAY
3 SEVERAL TIMES A WEEK
4 ONCE A WEEK
5 LESS THAN ONCE A WEEK
6 NEVER
0 NAP
8 DK
9 NA

PRAYER The United States Supreme Court has ruled that no state or local government may require the reading of the Lord's Prayer or Bible verses in public schools. What are your views on this--do you approve or disapprove of the court ruling?

1 APPROVE
2 DISAPPROVE
0 NAP

8 DK
9 NA

PRESTG80 Prestige of respondent's occupation (1980) Measure of respondent's occupational prestige which is a replication and extension of the Hodge-Siegel-Rossi ratings of occupational prestige. Low scores indicate low prestige and high scores indicate high prestige. See the appendices in the Cumulative Codebook for the General Social Surveys for more information.

0 DK, NA, NAP

PRESTIGE Prestige of respondent's occupation (1970) Measure of respondent's occupational prestige which is a replication and extension of the Hodge-Siegel-Rossi ratings of occupational prestige. Low scores indicate low prestige and high scores indicate high prestige. See the appendices in the Cumulative Codebook for the General Social Surveys for more information.

0 DK, NA, NAP

PRESTG1 Recoded prestige of respondent's occupation (see PRESTG80 and PRESTIGE). (Quartiles are approximate. Percentages will not always equal 25%.)

0 DK, NA, NAP
1 LOWEST QUARTILE
2 SECOND QUARTILE
3 THIRD QUARTILE
4 HIGHEST QUARTILE

RACE What race do you consider yourself?

1 WHITE
2 BLACK
3 OTHER

RACECEN1 RACECEN2 RACECEN3 What race do you consider yourself? Respondent’s first, second, and third mention.

1 WHITE
2 BLACK OR AFRICAN AMERICAN
3 AMERICAN INDIAN OR ALASKA NATIVE
4 ASIAN INDIAN
5 CHINESE
6 FILIPINO
7 JAPANESE
8 KOREAN

9 VIETNAMESE
10 OTHER ASIAN
11 NATIVE HAWAIIAN
12 GUAMANIAN OR CHAMORRO
13 SAMOAN
14 OTHER PACIFIC ISLANDER
15 SOME OTHER RACE
16 HISPANIC
98 DK
99 NA

RACLIVE Are there any (Negroes/Blacks/African-Americans) living in this neighborhood now? [see footnote 2]

1 YES
2 NO
0 NAP
8 DK
9 NA

RACOPEN Suppose there is a community-wide vote on the general housing issue. There are two possible laws to vote on. Which law would you vote for? [see footnote 2]

1 OWNER DECIDES
2 CAN'T DISCRIMINATE
0 NAP
8 DK
9 NA

REGION Region of interview

1 NEW ENGLAND
2 MIDDLE ATLANTIC
3 EAST NORTH CENTRAL
4 WEST NORTH CENTRAL
5 SOUTH ATLANTIC
6 EAST SOUTH CENTRAL
7 WEST SOUTH CENTRAL
8 MOUNTAIN
9 PACIFIC

RELIG What is your religious preference? Is it Protestant, Catholic, Jewish, some other religion, or no religion?

1 PROTESTANT
2 CATHOLIC
3 JEWISH
4 NONE
5 OTHER
8 DK
9 NA

RELGOSA Religiosity. Religiosity is the strength of an individual's attachment to his or her religious affiliation. It was created by combining three other variables- RELITEN (attachment to religious preference), PRAY (frequency of prayer), and ATTEND (frequency of church attendance). If an individual says he has a strong attachment to his religious affiliation, attends church often, and prays often, then he or she is highly religious. If an individual says he doesn't have a strong attachment to his religious affiliation, attends church infrequently, and prays infrequently, then he or she is not religious. Everyone else will be somewhere between highly religious and not religious. (Not available for 1982 and 1975.)

1 HIGH IN RELIGIOSITY
2 MEDIUM IN RELIGIOSITY
3 LOW IN RELIGIOSITY
9 MISSING

RELGOSB Religiosity. Religiosity is the strength of an individual's attachment to his or her religious affiliation. It was created by combining three other variables- RELITEN (attachment to religious preference) and ATTEND (frequency of church attendance). Note that PRAY was not used in this index. If an individual says he has a strong attachment to his religious affiliation and attends church often, then he or she is highly religious. If an individual says he doesn't have a strong attachment to his religious affiliation and attends church infrequently, then he or she is not religious. Everyone else will be somewhere between highly religious and not religious.

1 HIGH IN RELIGIOSITY
2 MEDIUM IN RELIGIOSITY
3 LOW IN RELIGIOSITY
9 MISSING

RELITEN Would you call yourself a strong (Protestant, Catholic, Jew, Other-specify) or a not very strong (Protestant, Catholic, Jew, Other-specify)?

1 STRONG
2 NOT VERY STRONG
3 SOMEWHAT STRONG
4 NO RELIGION
0 NAP

8 DK
9 NA

RINCOM98 Respondent's income. In which of these groups did your earnings for last year fall? That is, before taxes or other deductions [see footnote 3].

1 Under $10000
2 $10000-14999
3 $15000-19999
4 $20000-24999
5 $25000-34999
6 $35000-49999
7 $50000-59999
8 $60000-74999
9 $75000-89999
10 $90000-109999
11 $110000+

RINCOM91 Respondent's income. In which of these groups did your earnings for last year fall? That is, before taxes or other deductions [see footnote 3].

1 UNDER $1000
2 $10000-14999
3 $15000-19999
4 $20000-24999
5 $25000-34999
6 $35000-49999
7 $50000-59999
8 $60000-74999
9 $75000+

RINCOM86 Respondent's income. In which of these groups did your earnings for last year fall? That is, before taxes or other deductions [see footnote 3]

1 UNDER $1000
2 $10000-14999
3 $15000-19999
4 $20000-24999
5 $25000-34999
6 $35000-49999
7 $50000-59999
8 $60000+

RINCOM82 Respondent's income. In which of these groups did your earnings for last year fall? That is, before taxes or other deductions [see footnote 3]

1	UNDER $1000
2	$10000-14999
3	$15000-19999
4	$20000-24999
5	$25000-34999
6	$35000-49999
7	$50000+

RINCOME Respondent's income. In which of these groups did your earnings for last year fall? That is, before taxes or other deductions [see footnote 3]

1	UNDER $1000
2	$10000-14999
3	$15000-19999
4	$20000-24999
5	$25000+

SATFIN We are interested in how people are getting along financially these days. So far as you and your family are concerned, would you say that you are pretty well satisfied with your present financial situation, more or less satisfied, or not satisfied at all?

1	SATISFIED
2	MORE OR LESS
3	NOT AT ALL SATISFIED
8	DK
9	NA

SEI Respondent's socioeconomic index. Measure of occupational prestige which is based on the work of Otis Dudley Duncan. Low scores indicate low prestige and high scores indicate high prestige. See the appendices in the Cumulative Codebook for the General Social Surveys for more information. (Quartiles are approximate. Percentages will not always equal 25%.)

0	NAP

SEI1 Recoded respondent's socioeconomic index (see SEI)

0	NAP
1	LOWEST QUARTILE
2	SECOND QUARTILE
3	THIRD QUARTILE

4 HIGHEST QUARTILE

SEX Respondent's sex

1 MALE
2 FEMALE

SPKATH There are always some people whose ideas are considered bad or dangerous by other people. For instance, somebody who is against all churches and religion. If such a person wanted to make a speech in your (city/town/community) against churches and religion, should he be allowed to speak, or not?

1 ALLOW
2 NOT ALLOW
0 NAP
8 DK
9 NA

SPKCOM Now, I should like to ask you some questions about a man who admits he is a Communist. Suppose this admitted Communist wanted to make a speech in your community. Should he be allowed to speak, or not?

1 ALLOW
2 NOT ALLOW
0 NAP
8 DK
9 NA

SPKHOMO And what about a man who admits that he is a homosexual? Suppose this admitted homosexual wanted to make a speech in your community. Should he be allowed to speak, or not?

1 ALLOW
2 NOT ALLOW
0 NAP
8 DK
9 NA

SPKMIL Consider a person who advocates doing away with elections and letting the military run the country. If such a person wanted to make a speech in your community, should he be allowed to speak, or not?

1 ALLOW
2 NOT ALLOW
0 NAP

8 DK
9 NA

SPKRAC Or consider a person who believes that Blacks are genetically inferior. If such a person wanted to make a speech in your community claiming that Blacks are inferior, should he be allowed to speak, or not?

1 ALLOW
2 NOT ALLOW
0 NAP
8 DK
9 NA

SUICIDE1 Do you think a person has the right to end his or her own life if this person has an incurable disease?

1 YES
2 NO
0 NAP
8 DK
9 NA

SUICIDE2 Do you think a person has the right to end his or her own life if this person if this person has gone bankrupt?

1 YES
2 NO
0 NAP
8 DK
9 NA

SUICIDE3 Do you think a person has the right to end his or her own life if this person has dishonored his or her family?

1 YES
2 NO
0 NAP
8 DK
9 NA

SUICIDE4 Do you think a person has the right to end his or her own life if this person is tired of living and ready to die?

1 YES
2 NO

0 NAP
8 DK
9 NA

TOLATH Tolerance-anti-religionist. This is a measure of tolerance created by combining COLATH, LIBATH, and SPKATH. These questions ask if a person who is an atheist should be allowed to teach in a college or university, to have a book in the public library, and to give a public speech. A person is high in tolerance if they would allow all three and low if they would not allow any. All others are medium in tolerance. (Not available for 1975.)

3 HIGH
4 MEDIUM HIGH
5 MEDIUM LOW
6 LOW
9 MISSING

TOLCOM Tolerance-communist. This is a measure of tolerance created by combining COLCOM, LIBCOM, and SPKCOM. These questions ask if a person who is a Communist should be allowed to teach in a college or university, to have a book in the public library, and to give a public speech. A person is high in tolerance if they would allow all three and low if they would not allow any. All others are medium in tolerance. (Not available for 1975.)

3 HIGH
4 MEDIUM HIGH
5 MEDIUM LOW
6 LOW
9 MISSING

TOLHOMO Tolerance-homosexual. This is a measure of tolerance created by combining COLHOMO, LIBHOMO, and SPKHOMO. These questions ask if a person who is a homosexual should be allowed to teach in a college or university, to have a book in the public library, and to give a public speech. A person is high in tolerance if they would allow all three and low if they would not allow any. All others are medium in tolerance. (Not available for 1975.)

3 HIGH
4 MEDIUM HIGH
5 MEDIUM LOW
6 LOW
9 MISSING

TOLMIL Tolerance-militarist. This is a measure of tolerance created by combining COLMIL, LIBMIL, and SPKMIL. These questions ask if a person who is a militarist should be allowed to teach in a college or university, to have a book in the public library, and to give a public speech. A person is high in tolerance if they would allow

all three and low if they would not allow any. All others are medium in tolerance. (Not available for 1975.)

3	HIGH
4	MEDIUM HIGH
5	MEDIUM LOW
6	LOW
9	MISSING

TOLRAC Tolerance-racist. This is a measure of tolerance created by combining COLRAC, LIBRAC, and SPKRAC. These questions ask if a person who is a racist should be allowed to teach in a college or university, to have a book in the public library, and to give a public speech. A person is high in tolerance if they would allow all three and low if they would not allow any. All others are medium in tolerance. (Not available for 1975.)

3	HIGH
4	MEDIUM HIGH
5	MEDIUM LOW
6	LOW
9	MISSING

WEIGHT See Appendix B for a discussion of the weight variable.

YEAR Year of survey

Footnote 1 – INCOME98 was used for 2004 and 2002, INCOME91 was used for 1996, INCOME86 was used for 1989, INCOME82 was used for 1982, and INCOME was used for 1975. While the categories are different, they can be recoded to allow comparisons over time.

Footnote 2 -- In 1975, this question was asked of whites only. In 1982, 1989, 1996, 2002, and 2004, this question was asked of all races. This must be kept in mind when comparing data over time.

Footnote 3 – RINCOM98 was used for 2004 and 2002, RINCOM91 was used for 1996, RINCOM86 was used for 1989, RINCOM82 was used for 1982, and RINCOME was used for 1975. While the categories are different, they can be recoded to allow comparisons over time.

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Appendix B: Notes to the Instructor

DATA

The data for this module are available as SPSS for Windows portable files. There is one file containing only the 2004 sample data (called gss_04.por) and another file containing the variables in all time periods (called gss_04_02_96_89_82_75.por). These files can be downloaded from the Social Sciences Teaching Resources Depository website located at http://www.ssric.org.

SPSS FOR WINDOWS

There is an introduction to the basics of SPSS for Windows available for your use, *SPSS for Windows Version 13.0: A Basic Tutorial* (Linda Fiddler, Laura Hecht, Edward Nelson, Elizabeth Ness Nelson, James Ross). This can be ordered from McGraw-Hill (ISBN 0-07-353671-7). It is a basic introduction to SPSS and can be used as a supplement for a class or as a tutorial to learn SPSS by oneself. The version 13.0 book is currently on the web at http://www.csub.edu/~jross/projects/spss and will soon be moved over to the Social Sciences Teaching Resources Depository website at http://www.ssric.org.

CODEBOOK

The codebook is included in this module as Appendix A. A list of all the variables for student use is at the beginning of the codebook. A variable (YEAR) was created specifying the year of the survey. If you are using the file containing all time periods, remember that you must either select out the cases for one of the years or use YEAR as a control variable in a crosstabulation to be able to distinguish one year from the other year. For the exercises on cohort analysis, we suggest that you recode YEAR and combine the 2002 and 2004 surveys into one time period and treat that as if it was the 2003 survey.

Several of the variables have been recoded to make them easier to use. AGE was recoded into AGE1 and AGE2 to reduce the number of categories. EDUC was also recoded into EDUC1 for the same reason.

There are two occupational prestige variables (PRESTIGE and PRESTG80). PRESTIGE is based on a code introduced in 1970 and used in the 1975 and 1982 surveys. PRESTG80 is based on a code introduced in 1980 and used in the 1989, 1996, 2002, and 2004

surveys. These variables were recoded to make it possible to combine them into a single variable (PRESTG1).

Total family income has five variables. INCOME was used in the 1975 survey. INCOME82 is based on a code introduced in 1982 and was used in the 1982 survey. INCOME86 is based on a code introduced in 1986 and was used in the 1989 survey. INCOME91 is based on a code introduced in 1991 and was used in the 1996 survey. INCOME98 is based on a code introduced in 1998 and was used in the 2002 and 2004 surveys. We didn't attempt to combine these variables into a single variable. Respondent's income also has five variables (RINCOME, RINCOM82, RINCOM86, RINCOM91, RINCOM98) and follows the same pattern as total family income.

There is a socioeconomic status index (SEI) in the GSS that we recoded to make it easier to use (SEI1).

Race in the 1975, 1982, 1989, and 1996 surveys consisted of three categories (White, Black, Other). In 2002 and 2004 the GSS included three race variables (RACECEN1, RACECEN2, RACECEN3). These variables are the respondent's first three answers to the question "What race do you consider yourself to be?"

We also created two new variables to measure religiosity (RELIGOSA, RELIGOSB) and five new variables to measure tolerance (TOLATH, TOLCOM, TOLHOMO, TOLMIL, TOLRAC). See the codebook in Appendix A for a description of these new variables.

WEIGHTING

There is a weight variable (WEIGHT) in the data set that should be used. The population for the survey was adults (18+). If there was only one eligible person in the household that person was interviewed. But if there were two or more eligible people, then one of them was randomly selected. If there were two eligible people, then each person had a 1 out of 2 chance of being selected. If there were three eligible people, then each eligible person had a 1 out of 3 chance and so on. In other words, the more eligible people in the household, the smaller the chance of selection for any one of them. We can correct for this by weighting each case by the number of eligible people in the household. There is a variable called ADULTS which is the number of people 18 years of age or older in the household and this is, of course, also the number of eligible people in the household. The weighted number of cases is just the number of eligible adults multiplied by the number of cases. This means that each case with two eligible adults had a weight twice that of each case with one eligible adult, each case with three eligible adults has a weight three times that of each case with one eligible, and so on. The problem with this is that it also inflates the total number of cases. So we have to adjust the weights so the number of cases stays approximately constant.

In 2004, the General Social Survey modified their sampling design and included two weight variables that they hadn't included in previous surveys. For 2004, we used one of their weight variables in the data set. For a more complete description of the weight variables in 2004, see Appendix A in the *General Social Survey 1972-2004 Cumulative Codebook* (James A. Davis, Tom W. Smith and Peter V. Marsden, available through the Roper Center for Public Opinion).

CUSTOMIZING THE MODULE

The module can be integrated into your classes in many ways. One approach is to assign the introductory chapters to be read by the students and work through the simpler exercises in class. Ideally students should be given the opportunity to use the computer themselves and not just watch you do it. Later, exercises can be assigned as homework and then reviewed in class. Chapters 4 and 6 contain a culminating exercise in which students choose a problem, produce the appropriate tables and statistics, interpret the tables, and write a brief research report.

A second approach is for faculty to develop their own exercises to accompany the module. There is no reason that you should limit yourself to the exercises we developed. If you do develop your own exercises, be sure to test them before assigning them.

A third approach is for faculty to use the data set accompanying the module but not use the module itself. You would have to develop your own introductory material to make it relevant to your particular class.

The only statistics used in the module are percentages, chi square, Cramer's V, and Gamma. You could skip over all these statistics except percentages or you could introduce other statistical techniques.

Two other appendices are included for your possible use. Appendix C contains supplemental instructional materials. Appendix D contains a description of how to compute the measures of association used in this module--Cramer's V and Gamma.

Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Appendix C:
Supplemental Instructional Materials

Reminder on Reading Tables

- Read the title carefully. (A good title tells exactly what is in the table.)
- Look for headnotes or other explanations of the title.
- Look for any footnotes. Do they apply to the whole table? If not, to what part?
- Evaluate the source of the data. Is it likely to be reliable? Is it up to date?
- Study the column and row headings. What do they mean?
- Look for any overall average or total for the whole table. If the data are percentages, understand how they add to 100 percent.
- Describe some of the numbers in words. (For example, [Number] percent of [100%] are [row or column heading].) Sometimes population numbers are given in thousands, so you need to remember that that means millions.
- Look for patterns. What is the range? (For example, what are the largest and smallest percents?). Are most cases concentrated in one category, or are they more evenly spread out?
- Summarize the main points of the table in words, using figures from the table as examples.

Table C1 -- Marital Status* by Sex, March 1995
***Persons age 18 and over**

Marital Status	Women	Men
Married	59.2	62.7
Widowed	11.1	2.5
Divorced	10.3	8.0
Never Married	19.4	26.8
Total (percent)	100.0	100.0
Total (in thousands)	99,588	92,008

Source: Bureau of the Census. 1996. *Statistical Abstract of the United States:* 1996. Washington

DC: U. S. Government Printing Office. P.55

Sometimes it is not obvious how a table adds to 100%. (See Table C2.) A quick examination shows that it does not add to 100% down or across. (In 1890, 84% male + 18% female would be more than 100%.) To save space, this chart omits the percent of people who are not in the labor force. In 1890, the 84% of the males who were in the labor force plus the 16% of the males who were not in the labor force equals 100%. Similarly, 18% of the females were in the labor force and the rest of them were not in the labor force in 1890.

Table C2 -- Labor Force Participation Rate* by Sex: 1890-1995

*Percent of the noninstitutionalized population who are employed or who are unemployed and looked for work in the last four weeks. (Pre-1947 figures include those age 14 and over, later figures are age 16 and over).

Year	Male	Female
1890	84.3	18.2
1900	85.7	20.0
1920	84.6	22.7
1930	82.1	23.6
1940	82.5	27.9
1950	86.8	33.9
1955	86.2	35.7
1960	84.0	37.8
1965	81.5	39.3
1970	80.6	43.4
1975	78.4	46.4
1980	77.4	51.5
1985	76.3	54.5
1990	76.4	57.5
1995	75.0	58.9

Source: U.S. Bureau of the Census. 1976. *Historical Statistics of the United*

States: Colonial Times to 1970. (Bicentennial Ed.), Part I (1975), Washington DC: U..S. Government Printing Office, pp 131-32; U.S. Bureau of the Census. 1992. *Statistical Abstract of the United States: 1992*. Washington DC: U.S. Government Printing Office, p 383; U.S. Bureau of the Census. 1996. S*tatistical Abstract of the United States: 1996.*Washington DC: U.S. Government Printing Office, p.393.

Reminder and Exercises

Frequency and Percent Distributions, Figures and Graphs

Frequency and percent distributions use only a few numbers to describe data.

Frequency distributions use charts to list the range of possible values and the number of cases in each.

Percent distributions convert the frequencies to percentages. (Remember that the percent is the number per hundred, and the proportion is the part compared to the whole. You may want to review basic math--fractions, decimals, and percents, perhaps *Overcoming Math Anxiety* by Sheila Tobias.)

Exercise 1.a.

Construct a percent distribution using the information from public opinion polls on women's issues described in Chapter 2. Start with the percent who disapproved of a married woman working if her husband could support her.

Tables should have clear labels and definitions, so start your table with a title at the top that describes exactly what it is. Label the columns and rows carefully. Give the source of the data at the bottom of the table using one of the standard reference formats such as those recommended by the American Sociological Association or the American Psychological Association. (Note: Cite the source you used. Give the original source, e.g., Yankelovich, only if you actually looked it up and copied the figures from it. For this exercise, your source is Nelson, Elizabeth and Ed Nelson. 1997. *California Opinions on Women's Issues: 1985-1995.* Unpublished manuscript.

Figures and Graphs

We can present the same information visually as figures or graphs.

Bar graphs and Histograms use rectangles to show the number or percent in each interval. The intervals are marked along the horizontal axis (the bottom) and the frequencies or percents along the vertical axis (the left side), so zero for both scales is in the lower-left corner.

Histograms are used with ordered, discrete or continuous data. Since age in years can be ordered and is continuous from birth to old age, we would put the bars right next to each other.

Bar graphs use separate rectangles for each unit of nonordered, discrete data. For example, marital status--married or single--cannot be quantified or ordered, so we would use a separate bar for each category.

Frequency Polygons use dots at the midpoints of each interval in a similar way. Notice that it would make sense to use a frequency polygon with ordered data such as age but not for nonordered data such as race.

Exercise 1.b.

Construct a histogram or bar chart using the same data as exercise 1.a. Remember that frequencies or percents are usually marked along the left side of the chart with the smallest numbers at the bottom and the values on the base start with the lowest values and go from left to right so both scales use the same zero.

The overall impression produced by a graph depends on the ratio of the measurements of the horizontal and vertical scales. It is important to communicate these quantitative relationships accurately. Experiment with different scales on scratch paper to find one that seems to be a useful way to illustrate your data. Connect the midpoints of the bars so you can see a frequency polygon. Again, be sure to use clear labels. Usually the title of a figure or graph (called the legend) is on the bottom. Include the source of your data in American Sociological Association or American Psychological Association style.

Example with Frequency Distribution, Percent Distribution, Dummy Table, and Crosstab

Frequency and percent distributions use a few numbers to describe the data. **Frequency distributions** use charts to list the range of possible values and the number of cases in each category. **Percent distributions** are similar but convert the numbers to percentages. (Remember that the percent is the number per hundred, and the proportion is the part compared to the whole. You may want to review

basic math--fractions, decimals, and percents, perhaps *Overcoming Math Anxiety* by Sheila Tobias.) With the percent distribution, we can make comparisons even when the actual numbers are different.

This example comes from a study of opinions and behavior of California State University Students in 1994. The question related to the students' knowledge of pregnancy facts.

Table C3 -- Frequency and Percent Distribution of CSU Students' Responses to "At What Time in Her Monthly Cycle is a Woman Most Likely to become Pregnant?"

	Number	**Percent**
Beginning	545	27.0
Middle	1,113	55.1
End	361	17.9
Total	**2,019**	**100.0**

Source: Survey of California State University Students conducted by James Ross (1994)

The percentages show that over half (55%) of the CSU students answered the question correctly. (A woman is most fertile in the middle of the monthly cycle.)

Crosstabulation

To **analyze** means to break something down into its component parts and study them in order to gain a better understanding of the whole. We can look at these responses in more detail to gain a better understanding of students' knowledge of this rather important part of human life. **Crosstabulation** uses tables showing the number and percentage of cases in each combination of categories of the data. We might expect that students would have better understanding of health information related to their own bodies, so female students might be more likely than the male students to answer correctly. So, we hypothesize that females will be more likely to answer "middle." We can make a dummy table showing what we would expect is the hypothesis were supported by the data.

Dummy Table C4 -- Frequency and Percent Distribution of CSU Students' Responses to "At What Time in Her Monthly Cycle is a Woman Most Likely to Become Pregnant?"

	Male		Female
Beginning	a	>	b
Middle	c	<	d
End	e	>	f

The next table is a crosstabulation of the responses by sex, so we can look at similarities and differences in the responses of males and females. To crosstabulate by sex, we construct separate frequency and percent distributions for each sex, calculating the percentages down so we can compare across.

Table C5 -- Frequency and Percent Distribution of CSU Students' Responses to "At What Time in Her Monthly Cycle is a Woman Most Likely to Become Pregnant?" by Sex

	Male		**Female**		**Total**	
	Number	Percent	Number	Percent	Number	Percent
Beginning	279	30.8	266	23.9	545	27.0
Middle	448	49.4	665	59.7	1,113	55.1
End	179	19.8	182	16.4	361	17.9
Total	906	100.0	1,113	100.0	2,019	100.0

Source: Survey of California State University students conducted by James Ross (1994)

Women students were more likely than male students to answer correctly (60% of the women compared to 50% of the men answered that a woman is most likely to become pregnant in the middle of her monthly cycle).

SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Appendix D:
Computation of Measures of Association

There are many measures of association used to measure the strength of relationship. Each has advantages and disadvantages. In this module, we have used two--Cramer's V and Goodman and Kruskal's Gamma. This appendix describes how to compute these measures. However, you will use a statistical package such as SPSS for Windows to do the actual computations for the exercises.

Cramer's V is one of several measures based on chi square. Chi square itself is not a measure of association, but a test of the hypothesis that two variables are unrelated. V is equal to the square root of the following value--chi square divided by the product of the number of cases in the table and the smaller of two values--the number of rows minus one and the number of columns minus one. The chi square for table 3.2 was equal to 12.52. V would be equal to 12.52 divided by the square root of the product of 900 and 1 or 12.52 divided by 30 or 0.42.

Cramer's V should be used when one or both of the variables consist of an unordered set of categories. It varies from 0 to 1. The closer it is to 0, the weaker the relationship and the closer to 1, the stronger the relationship. V can never be negative.

Gamma assumes that both of the variables consist of ordered categories. To understand the computation of Gamma, you must think of pairs of cases. Imagine four individuals. We'll just call them A, B, C, and D. For each person, we know their income and education which has been categorized as low, medium, and high. The table below displays these values.

Education of Respondent			
Income	High	Medium	Low
High	A		
Medium		C	

Low	B		D

These four individuals form six possible pairs. A can be paired with B, A with C, A with D, B with C, B with D, and C with D. Notice that the AB pair is the same as the BA pair since they involve the same two individuals.

A has more education and income than C. This is what we call a **concordant pair**. A is higher on both variables (i.e., income and education) than C. The AD pair is also concordant. A is higher on both education and income than D. And the CD pair is also concordant. C has more education and income than D.

However, C has more income than B, but C has less education than B. This is what we call a **discordant pair**. C is higher on one variable (i.e., income), but lower on the other variable (i.e., education).

When we look at the AB pair we see another possibility. A has more income than B, but A and B are tied on education. This is what we call a tied pair. The BD pair is also tied. B has more education than D, but B and D have the same income.

Gamma ignores all tied pairs. Since two of the six possible pairs are tied, Gamma would be based on the remaining four untied pairs. Gamma is equal to the number of concordant pairs (C) minus the number of discordant pairs (D) divided by the sum of the number of concordant pairs and discordant pairs.

In this example, Gamma would equal (3-1)/(3+1) or 2/4 or 0.50. Since there are more concordant pairs than discordant pairs, we can observe that it is more common for pairs to have the same order on both variables than to have different orders. In other words, large amounts of education tend to go with large amounts of income and small amounts of education tend to go with small amounts of income. This is what we call a **positive relationship**. Gamma has a positive sign if the relationship is positive [Note 1].

If the number of discordant pairs had been greater than the number of concordant pairs, the relationship would have been **negative** and the sign of Gamma would have been negative. This would have meant that large values of one variable would tend to go with small values of the other variable and that small values of one variable would tend to go with large values of the other variable.

The numerical value of Gamma tells us the strength of the relationship. The closer the value of Gamma to 0, the weaker the relationship and the closer to 1, the stronger the relationship [Note 2].

Notes

1. This will be true only if the columns are arranged from high to low (left to right) and the rows are arranged from high to low (top to bottom).

2. The fact that Gamma ignores all tied pairs tends to inflate the value of Gamma. For this reason, Gamma produces a larger measure of association than other measures.

REFERENCES AND SUGGESTED READING

- Herman J. Loether and Donald G. McTavish, 1993, *Descriptive and Inferential Statistics An Introduction* (4th edition), Allyn and Bacon.
- Marija J. Norusis, 2005, *SPSS 13.0 Guide to Data Analysis*, Prentice Hall.
