

Author(s): Leege, David C., Lyman A. Kellstedt, and Kenneth D. Wald

Title: Religion and Politics: A Report on Measures of Religiosity in the 1989 NES Pilot Study

Date: February 1990

DataSet(s): 1988 National Election Study, 1989 Pilot Study

Keywords: Participation, Political Information, Religion

Abstract

Leege, Kellstedt, and Wald discuss the 1989 Pilot Study experimental measures of religiosity. The authors find that the branching scheme used to measure denominational affiliation in the Pilot Study seems to make interviewers more sensitive to distinctions within religious sects, resulting in more accurate religious classifications than traditional measures. In addition, the authors' analysis indicates that the experimental measures uncover a connection between religion and partisanship that is disguised by the imprecision of the traditional NES measure. Leege, Kellstedt, and Wald also examine the Pilot Study and 1988 NES items designed to assess a respondent's level of commitment to organized religion. They find that the traditional measures overstate attachment to religious institutions. Moreover, the additional Pilot Study questions allow for much greater sensitivity of measurement of the relationship between religion and political participation. The Pilot Study also included measures of cue-giving by denominational leaders. Leege, Kellstedt and Wald find that perceptions of cue giving differ by issue, denominational family, and frequency of presence at religious services. The information obtained through the Pilot Study, however, is not complex enough to determine whether cue giving affects the political attitudes of religious adherents. The authors recommend retaining the cue-giving sequence, but argue that measures of cue direction should be added. Finally, Leege, Kellstedt, and Wald examine and rank a number of other items relating to religious exposure and participation, such as measures of church attendance and self-identification as a "Born Again" Christian. The authors prepared a supplemental report, which provides further support for adopting the experimental Pilot Study measures of religious preferences. Leege, Kellstedt, and Wald find that the new codes and "religious family" designations more accurately reflect modern religious affiliation patterns. These more sensitive measures can better uncover different patterns of political behavior, both within and across denominations. Moreover, the Pilot Study filters reduce error in the measurement of the proportion of people who claim a religious preference. The authors conclude by proposing specific question formats and coding schemes for the 1990 National Election Study.

**The Impact of Verbal Labeling of Response Alternatives and Branching
on Attitude Measurement Reliability in Surveys**

Jon A. Krosnick

and

Matthew K. Berent

The Ohio State University

February, 1990

The authors wish to express their gratitude to Konnie Toth, Schelli Cavanaugh, Peggy Caldwell, and Tom Herby for conducting the interviewing for Study 1, to Sondra Crayton and Regina Owens for conducting the interviewing for Study 4, and to David Boninger and Stanley Presser for helpful comments. Correspondence should be addressed to Jon A. Krosnick, Department of Psychology, 108 Lazenby Hall, 1827 Neil Avenue, The Ohio State University, Columbus, Ohio, 43210-1222.

**The Impact of Verbal Labeling of Response Alternatives and Branching
on Attitude Measurement Reliability in Surveys**

Abstract

This study assessed whether verbal labeling of response alternatives and branching formats improve the reliability of survey attitude measures. Respondents in two telephone surveys, one self-administered survey, and one face-to-face survey were interviewed on two occasions separated by between one and three months and were randomly assigned to receive either fully-labeled branching or partially-labeled non-branching attitude measures during both interviews. As expected, the reliability of the fully-labeled branching questions was greater than that of the partially-labeled non-branching questions. It appears that this difference is fully attributable to the effect of the branching format rather than to the verbal labels *per se*. This suggests that survey researchers should employ branching formats in attitude questions whenever possible in order to maximize reliability. These findings also contribute to a growing literature calling into question the widely-held belief in political science that political party identification is more persistent over time and more psychologically consequential than citizens' attitudes toward government policies.

**The Impact of Verbal Labeling of Response Alternatives and Branching
on Attitude Measurement Reliability in Surveys**

Introduction

Since the earliest empirical research on political attitudes, political scientists have believed that political party identification is much more stable over time than citizens' attitudes toward government policies. This wide-spread belief is based almost exclusively on evidence from the 1956-1960 and 1972-1976 National Election Panel Studies (NES), surveys that tracked changes in Americans' political attitudes over four-year periods including two presidential election campaigns (Converse, 1964; Converse & Markus, 1979). More recently, the same conclusion has been supported by analyses of the 1980 National Election Panel Study, which tracked political attitude change over a nine-month period during a single presidential election campaign (Markus, 1982). On the basis of this evidence, political scientists often make claims such as "partisan loyalties are more stable over time than any other political attitude (Abramson, 1983, p. 99)."

Although this conclusion is certainly consistent with the accumulated evidence from the NES panel surveys, there is an alternative interpretation for these findings. In all previous studies, attitude stability was assessed by examining zero-order test-retest correlations, and it is well-known that these correlations reflect both the amount of attitude change that takes place during a given time period and the amount of random measurement error in reports of the attitude (see, e.g., Alwin, 1973). Therefore, evidence that reports of party identification are more consistent over time than are reports of policy

attitudes may be due to differences in either or both of these processes.

In fact, there is good reason to suspect that the NES party identification measures may be more reliable than the policy attitude measures in those surveys. Party identification has been measured via a series of branching questions in which respondents are first asked whether they consider themselves to be Republicans, Democrats, or independents. People who report identifying with a party are then asked whether they do so strongly or weakly, and people who say they are independents are asked whether they lean toward one party or the other. As a result, respondents were segmented into seven groups along a continuum ranging from strong Republican to strong Democrat. The verbal labels on the response alternatives presumably make it easy for respondents to understand their meanings. Consequently, respondents' choices are likely to be highly reliable.

In contrast, the response alternatives offered by the NES policy attitude measures appear to be markedly more ambiguous. These measures typically asked respondents to place themselves on 7-point scales on which only the end points are labeled with words. For example, the 1972 measure of attitudes toward guaranteed full employment read as follows: "Some people feel that the government in Washington should see to it that every person has a job and a good standard of living. Suppose that these people are at one end of this scale -- at point number 1. Others think the government should just let each person get ahead on his own. Suppose that these people are at the other end -- at point number 7. And, of course, some other people have opinions in between. Where would you place yourself on this scale, or haven't you thought much about

this?" If the lack of labeling increases the ambiguity of the meanings of the mid-scale response alternatives, it may decrease the reliability of reports, especially those made by respondents with moderate attitudes. The difference in verbal labeling could therefore at least partly account for the reduced over-time consistency of policy attitude reports as compared to reports of party identification.

The NES party identification and policy attitude measures differ not only in the extent of verbal labeling of response alternatives but also because the party identification measure involves branching, whereas the policy attitude measures do not. That is, the party identification measure first asks respondents to report the direction of their attitudes, and then asks them to report the intensity of those attitudes. In contrast, the policy attitude measures ask respondents to indicate attitude direction and intensity in a single response. Armstrong, Denniston, and Gordon (1975) have shown that decomposing a complex decision task into simpler component decisions increases the precision of the final result. It may therefore be that decomposing attitude reports through branching may increase their reliability. This could also partly explain why party identification reports are more consistent over time than policy attitude reports in the NESs.

A variety of studies offer indirect support for these propositions. First, a number of experimental comparisons of fully-labeled and partially-labeled attitude measures indicate that the former are more reliable than the latter (Madden, 1960; Peters & McCormick, 1966; Zaller, 1988; c.f. Finn, 1972). Second, in an analysis of the NES panel data that decomposed test-retest correlations into components due to attitude stability and attitude measurement unreliability, Krosnick and

Alwin (1989a) found that the NES party identification measure is more reliable than the NES policy attitude measures. However, party identification was no more stable than policy preferences; both party identification and policy attitudes were highly stable.

Although Krosnick and Alwin's (1989a) findings are consistent with the claim that item format differences are responsible for the misleading appearance that party identification is more stable than policy attitudes, there are other possible explanations for their finding. The amount of random error in attitude reports is a function of the ambiguity of response alternatives, the ambiguity of respondents' internal attitudinal cues, and the prevalence of non-attitudes (for an extensive discussion of these issues, see Alwin & Krosnick, 1989). According to a great deal of research in psychology, people are able to report some attitudes quickly and confidently, due to clear and accessible internal cues, whereas other attitudes are reported more slowly, effortfully, and less confidently because relevant internal cues are more ambiguous and/or less accessible (Fazio, 1986; Fazio, Chen, McDonel, & Sherman, 1982; Krosnick, 1989). Similarly, social judgment theory (Sherif & Hovland, 1953, 1961; Sherif, Sherif, & Nebergall, 1965) suggests that people are often aware that their attitude toward an object falls within a particular range (called their latitude of acceptance) on an attitude continuum but may be uncertain about exactly which point within that range best represents their attitude. This ambiguity in internal attitudinal cues may contribute to unreliability, so policy attitude reports may be more unreliable than reports of party identification because the former are associated with more ambiguous cues.

A second alternative source of unreliability is non-attitudes. In 1964, Philip Converse argued that for any given survey attitude measure, respondents can be grouped into two classes: those who hold a pre-existing attitude toward the object and those who do not. Converse claimed that because respondents in the latter category want to appear opinionated to their interviewers, they sometimes choose randomly from among the response alternatives offered, thus reporting what he called non-attitudes. If survey respondents do sometimes report non-attitudes, and if these reports are indeed purely random, they would enhance the apparent unreliability of an attitude measure. Therefore, party identification may appear to be more reliably measured than policy attitudes because the NES policy attitudes attract more reports of non-attitudes than does the NES party identification measure.

In order to determine if verbal labeling of response alternatives and branching affect the reliability of survey attitude measures, we conducted four experiments. In each, a series of political attitudes were assessed either using fully verbally labeled/branching questions or using non-branching questions involving scales with only the end points being verbally labeled. Respondents were reinterviewed between one and three months later, and the reliabilities of the two question forms were compared by examining the over-time consistency of respondents' reports.

STUDY 1

Method

Sample

For our first study, telephone interviews were conducted with residents of the Columbus, Ohio, metropolitan area. Telephone numbers were generated by a random digit dialing (RDD) method. Of the 167

eligible respondents who were contacted, 69 (or 41%) agreed to participate in initial interviews.¹ Approximately one month later, we successfully recontacted 63 (or 91%) of these individuals for a second interview.

Data Collection

Four trained telephone interviewers were randomly assigned to telephone numbers for initial interviews, and the interviewers were randomly assigned to respondents to conduct the follow-up interviews. At the end of the initial interview, interviewers asked respondents whether they would be willing to be contacted in a month so the investigators could study how their attitudes change over time. All of the respondents who completed an initial interview consented to a second interview and told the interviewer their first name.

The interviewers were not informed about the hypotheses being tested in the study. When probed after the interviews were completed, none of the interviewers were able to guess the hypotheses.

Measures

All respondents were asked about their political party identification, their ideological orientation, and their attitudes toward federal aid to minority groups, defense spending, and U.S. involvement in Central America. Two sets of questions measuring these attitudes were constructed for this experiment, fully-labeled/branching versions and partially-labeled versions. The fully-labeled/branching versions of the party identification and ideology measures are nearly identical to the NES measures of these variables, and the partially-labeled versions of the policy attitude measures are nearly identical to the NES measures of these variables. The fully-labeled/branching policy

attitude questions were written to parallel the structure of the fully-labeled/branching NES party and ideology questions while retaining the meanings of the NES policy questions as much as possible. Similarly, the partially-labeled party identification and ideology questions were written to parallel the structure of the partially-labeled NES policy attitude measures while retaining the meanings of the NES party and ideology questions as much as possible (for exact question wordings, see the Appendix). Each respondent was randomly assigned to receive either the fully-labeled/branching questions or the partially labeled questions during both interviews.

Assessment of Reliability

There are a variety of ways to assess the reliability of attitude measures, including test-retest correlations and unstandardized regression coefficients. Estimates generated by these methods are affected by the variances and distributions of the variables involved and can therefore sometimes be misleading. Therefore, for this study, we chose a very simple method to assess reliability: the percentage of respondents who provided the same response to a question at both times. The larger this percentage, the greater consistency there is in responses, and the more reliable the measure presumably is.

Results

Table 1 displays estimates of the reliabilities of the five attitude reports separately for the two question forms. The first point to note here is that the traditional NES party identification measure, the fully-labeled/branching version, is markedly more reliable (69.4%) than the traditional NES policy attitude measures, the partially-labeled versions (33.3%, 51.9%, and 29.6% for aid to minorities, defense

spending, and Central America, respectively). The traditional party identification reports were statistically significantly more reliable than reports of attitudes regarding Central America ($\chi^2(1)=9.81$, $p<.002$) and regarding aid to minorities ($\chi^2(1)=8.10$, $p<.005$), though only marginally significantly more reliable than reports of attitudes regarding defense spending ($\chi^2(1)=2.02$, $p=.15$). In general, then, these results support the assertion that the traditional NES party identification measure is more reliable than the traditional NES policy attitude measures. These findings are also consistent with Krosnick & Alwin's (1989a) conclusion in this regard based on structural equation analyses of the 1970s and 1980s NES panel data sets.

As expected, the combined reliability of the five fully-labeled/branching questions (61.9%) is substantially and significantly greater than the combined reliability of the five partially-labeled questions (38.5%, $\chi^2(1)=17.01$, $p<.001$). The largest gain in reliability occurred in the case of party identification, (39.8%, $\chi^2(1)=9.81$, $p<.002$). The policy attitude items showed somewhat smaller but nonetheless substantial gains (Minorities: 25.0%, $\chi^2(1)=3.87$, $p<.05$; Defense Spending: 25.9%, $\chi^2(1)=4.67$, $p<.04$; Central American : 23.2%, $\chi^2(1)=3.38$, $p<.07$). And, surprisingly, the reliability of ideological orientation reports was apparently unaltered by the verbal labeling (3.3%, $\chi^2(1)=0.07$, ns). In general, then, these results indicate that the fully-labeled/branching questions were more reliable than the partially-labeled questions.

We expected that the smallest gains in reliability due to verbal labeling/branching would occur for respondents with extreme attitudes. Because the most extreme response alternatives were verbally labeled on

both the fully-labeled/branching and the partially-labeled questions, the response alternatives of most interest to individuals with extreme attitudes were equally unambiguous on both forms. In contrast, the response alternatives of most interest to individuals with moderate attitudes varies a great deal in terms of their ambiguity between the fully-labeled/branching and partially-labeled forms. Therefore, these individuals would be expected to show the greatest gains in reliability due to full verbal labeling.

In order to test this hypothesis, we conducted two separate analyses, one using the extremity of attitude reports at wave one, and the second using the extremity of attitude reports at wave two. As the figures in Table 2 illustrate, both ways of doing the analysis confirmed our expectations. Respondents with extreme attitudes at either wave one or wave two were highly likely to provide the same attitude report at both waves, and the verbal labeling/branching did not significantly alter this likelihood. In contrast, respondents with moderate attitudes were much more likely to provide the same response at both waves if asked the fully-labeled/branching questions than if asked the partially-labeled questions.

Among respondents who placed themselves at 2 or 6 during the first interview, 62% of those asked the fully-labeled/branching questions placed themselves at the same point during the second interview, whereas only 28% of those asked the partially-labeled questions placed themselves at the same point during the second interview ($\chi^2(1)=7.71$, $p<.01$; see row 2 of Table 2). The same result holds for respondents who placed themselves at either 2 or 6 during the second interview, and for those who placed themselves at either 3 or 5 during the first or second

interview. The same trend appears for respondents who placed themselves at 4 during either interview, but this trend is only statistically significant for respondents who placed themselves at 4 during the second interview (see row 8 or Table 2). Given that the trend appears in row 4 of Table 2 as well, we are inclined to treat it as real in both cases. But among respondents who placed themselves at 1 or 7 during either interview, the addition of verbal labels did not enhance the reliability of their responses. Therefore, these results support the contention that the verbal labeling/branching manipulation had more impact on respondents with moderate attitudes than on respondents with extreme attitudes.

We next turned to the question of whether, once item format is controlled, reports of party identification are just as reliable as reports of policy attitudes. As expected, the figures in the first column of Table 1 indicate that party identification was equally reliably or less reliably measured than were the policy attitudes. Party identification was not significantly more reliable than aid to minorities attitude reports ($\chi^2(1)=0.09$, ns), Central America attitude reports ($\chi^2(1)=0.00$, ns), or a combination of the three policy attitudes ($\chi^2(1)=0.66$, ns). The only marginally significant difference here is that party identification was less reliably measured than defense spending attitudes ($\chi^2(1)=2.76$, $p<.10$).

The story is essentially the same according to the figures in column 2 of Table 1. Although the reliability of party identification is significantly greater than that of Central America attitudes ($\chi^2(1)=12.51$, $p<.001$), the reliability of party identification reports is not significantly different from that of defense spending attitudes

($\chi^2(1)=0.64$, ns), aid to minorities attitudes ($\chi^2(1)=0.96$, ns), or a combination of the three policy attitudes ($\chi^2(1)=2.44$, ns). In general, then, these results suggest that reports of party identification and policy attitudes are equally reliable when the formal properties of the measures are held constant.

STUDY 2

The results of Study 1 are generally consistent with our expectations. However, the sample of respondents who provided the data for that study were not necessarily comparable to more heterogeneous samples typical of national surveys. We therefore attempted to replicate the findings of Study 1 using a representative national sample.

In this study, we were also able to examine the impact on reliability of an additional formal property of attitude questions: whether or not a "don't know" filter is included. In most NES studies, the policy attitude and ideology questions have offered respondents the option to say that they had not thought about an issue before, whereas the party identification question did not. These filters typically attract between 10% and 30% of respondents when offered. Though unlikely, it is conceivable that removing these respondents from a sample lowers the reliability of an item (see Alwin & Krosnick, 1989). Thus, another potential formal source of non-comparability across these questions that may be responsible for reliability differences is the presence or absence of a "don't know" filter. We explored this issue in the present study.

Finally, this study allowed us to explore an additional issue of interest to public opinion researchers. We noted above that Converse

(1964) suggested non-attitudes may be one source of random error in attitude reports. It is widely believed that including "don't know" filters encourages respondents with no attitude on an issue to say so rather than flipping a mental coin and randomly choosing from among the response alternatives. This reasoning would suggest that including "don't know" filters should remove respondents who would provide unreliable answers from the sample and should therefore yield more reliable reports. This prediction, though widely taken for granted by survey researcher, has not yet been tested experimentally. And although one correlational analysis found support for this assumption (Andrews, 1984), another did not (Alwin & Krosnick, 1989). This study provided us with an opportunity to test this prediction experimentally.

Method

Sample

The data for this study were collected as a part of the 1989 National Election Pilot Study. A total of 1640 individuals constituting a representation national sample of American adults had been interviewed twice the year before for the 1988 NES and had provided their telephone numbers to interviewers. A stratified random sample of 855 of these individuals were selected to be reinterviewed for the 1989 NES Pilot. In fact, 614 individuals were successfully interviewed for the first wave during July and August, 1989, and 494 of them were successfully reinterviewed during September and October. Thus, the response rate for wave one was 74%; the response rate for wave two was 80%; and the time lag between interviews varied from one month to three months.

Data Collection

For both waves, respondents were interviewed by telephone by trained interviewers on the staff of the University of Michigan's Survey Research Center.

Measures

In the midst of long interviews addressing a wide range of political topics, all respondents answered questions about their political party identification, ideological orientation, and attitudes on defense spending, U.S. involvement in Central America, and gun control. Four sets of questions were constructed for this study. Two of the sets included partially-labeled response alternatives, and two included fully-labeled/branching response alternatives. One of the partially-labeled sets included a "don't know" filter in each question, and the other partially-labeled set did not. Similarly, one of the fully-labeled/branching sets included a "don't know" filter in each question, and the other did not. The texts of the questions are listed in the Appendix.

In contrast to the other questions, the ideology measure included a "don't know" filter in all of the question sets. Instead of varying the presence or absence of such a filter, the wording of the verbal labels was varied. In two question sets, the most extreme response alternatives were labeled "very" liberal or conservative. In the other two question sets, the more extreme response alternatives were labeled "extremely" liberal or conservative. This variation had no impact in the analyses reported below, so we do not discuss it further.

Each respondent was randomly assigned to receive one of the four question forms during both interviews.

Assessment of Reliability

Reliability was again assessed by the percentage of respondents who provided the same answer to each question during both interviews.

Results

Effect of Labeling/Branching

Table 3 displays reliability estimates for the partially-labeled and fully-labeled/branching formats. As expected, the reliability of the traditional NES party identification measure (67.2%) was again larger than the reliability of the traditional NES policy attitude measures addressing defense spending (40.4%, $\chi^2(1)=34.01$, $p<.0001$), Central America (32.2%, $\chi^2(1)=58.72$, $p<.0001$), gun control (45.6%, $\chi^2(1)=22.98$, $p<.0001$). Thus, the traditional measure of party identification was again more reliable than traditional measures of policy attitudes.

Also as expected, verbal labeling and branching increased reliability substantially. The combined reliability of the fully-labeled/branching items (56.6%) was significantly larger than the combined reliability of the partially-labeled items (44.3%, $\chi^2(1)=33.39$, $p<.00001$). The greatest gains in reliability occurred in the cases of Central America (16.6%, $\chi^2(1)=12.40$, $p<.001$) and gun control (16.5%, $\chi^2(1)=12.58$, $p<.001$). The gain in reliability was also significant in the case of defense spending (10.6%, $\chi^2(1)=4.84$, $p<.03$) but only marginally significant in the cases of party identification (7.8%, $\chi^2(1)=3.19$, $p<.08$) and ideological orientation (9.6%, $\chi^2(1)=3.64$, $p<.06$). The effect of verbal labeling was equally strong when a "don't know" filter was included and when it was not ($\chi^2(1)=2.58$, $p>.10$).²

Because the 1989 NES asked respondents about their educational attainment, it was possible to examine whether the effect of verbal labeling/branching varied across different educational levels (for a rationale, see Krosnick, 1990). To do so, respondents without a high school diploma were placed in the low education group, respondents with a high school diploma but no additional education were placed in the medium education group, and respondents with some education beyond high school were placed in the high education group.

As Table 4 illustrates, the fully-labeled/branching questions were more reliable than the partially-labeled questions for all three education groups. The difference was statistically significant for the low education (12.6%, $\chi^2(1)=4.94$, $p<.05$), medium education (17.2%, $\chi^2(1)=22.51$, $p<.001$), and high education groups (8.4%, $\chi^2(1)=7.81$, $p<.01$). An omnibus test of the interaction between education and question format was not statistically significant ($\chi^2(2)=3.51$, $p<.18$), which suggests that the effect of question format did not vary across education levels.

However, Rosnow and Rosenthal (1989) argued strongly against over-reliance on such omnibus tests and strongly advocated placing more weight on theoretically-based planned contrasts. Such comparisons between the education groups revealed that the effect of question format was not significantly different in the low and medium education groups ($\chi^2(1)=0.43$, ns), but it was marginally significantly stronger in the medium education group than in the high education group ($\chi^2(1)=3.50$, $p<.07$). This suggests that the effect of verbal labeling/branching on reliability may have been greatest for respondents with moderate levels of education.

In contrast to the results of Study 1, controlling for question format did not eliminate the difference in reliability between measures of policy attitudes and party identification. In the case of the partially-labeled items, party identification (59.4%) was significantly more reliable than defense spending (40.4%, $\chi^2(1)=17.16$, $p<.0001$), Central America (32.2%, $\chi^2(1)=36.04$, $p<.0001$), gun control (45.6%, $\chi^2(1)=9.40$, $p<.003$), and the combined reliability of the three policy attitude measures (39.5%, $\chi^2(1)=29.41$, $p<.0001$). In the case of the fully-labeled/branching questions, party identification (67.2%) was also significantly more reliable than defense spending (51.0%, $\chi^2(1)=12.37$, $p<.001$), Central America (48.8%, $\chi^2(1)=15.59$, $p<.0001$), and the combined reliability estimate of the three policy issue attitude measures (54.1%, $\chi^2(1)=12.48$, $p<.001$), although not from gun control (62.1%, $\chi^2(1)=1.36$, ns). These results therefore suggest that equating the extent of verbal labeling eliminates some but not all of the reliability discrepancy between party identification and policy attitudes.

"Don't Know" Filters

Consistent with many prior studies (Bishop, Oldendick, & Tuchfarber, 1980, 1983; Bishop, Oldendick, Tuchfarber, & Bennett, 1979; Schuman & Presser, 1981), including a "don't know" filter substantially increased the proportion of respondents who said they had no attitude toward each object (see Table 5). A larger percentage of respondents indicated "don't know" when a filter was present in the cases of party identification (11.2%, $\chi^2(1)=38.86$, $p<.0001$), defense spending (10.4%, $\chi^2(1)=25.80$, $p<.0001$), Central America (9.1%, $\chi^2(1)=19.23$, $p<.0001$), and the four items combined (8.2%, $\chi^2(1)=71.98$, $p<.0001$), though the

difference was not significant in the case of gun control (4.0%, $\chi^2(1)=1.63$, ns).

However, including a "don't know" filter did not decrease the amount of random error in attitude reports (see Table 6). The difference between the filtered and unfiltered forms was not statistically significant for party identification, (3.0%, $\chi^2(1)=0.47$, ns), defense spending (0.8%, $\chi^2(1)=0.03$, ns), Central America (4.0%, $\chi^2(1)=0.71$, ns), gun control (4.5%, $\chi^2(1)=0.93$, ns), or the combined reliability of the four items (3.7%, $\chi^2(1)=2.41$, ns). This finding clearly challenges the conventional wisdom that many respondents who are attracted by "don't know" filters would have offered random responses had the filter not been offered.

STUDY 3

In order to explore the generality of the effect of verbal labeling/branching observed in Studies 1 and 2, we conducted a third study. Whereas these studies involved telephone interviewing, Study 3 involved self-administered questionnaires. Utilizing this administration mode allowed us to test an alternative explanation for the findings of Studies 1 and 2. Because those studies both involved telephone interviewing, respondents had to hold the response alternatives to each question in memory before making their choice. Doing so is presumably relatively easy with the fully-labeled/branching questions; respondents must hold in memory at most three verbally-labeled alternatives at once. In contrast, the partially-labeled questions required that respondents hold a seven-point scale and the verbal labeled attached to the end points in memory. It is conceivable that some respondents receiving that format occasionally forgot which

verbal label was attached to which end-point, and this confusion produced random error in responses. By using self-administered questionnaires, we eliminate the need for respondents to hold response alternatives in memory, so we eliminate the impact of confusion on reliability.

In Study 3, we also explored the validity of a second possible alternative explanation for the findings of Studies 1 and 2. In those studies, reliability was assessed by the consistency of attitude reports on two occasions separated by between one and three months. Though unlikely, it is conceivable that during the second interviews, some respondents were able to recall their attitude reports from the first interview and simply stated those recollections. If people asked the fully-labeled/branching questions were able to remember their initial responses better, the apparently enhanced reliability of these questions may actually reflect better recall by respondents. This seems especially unlikely in Study 2, in which the questions of present interest were embedded in the midst of numerous questions on a variety of political topics. Nonetheless, we explored this possibility in this study.

Method

Sample

Seventy-seven undergraduates at The Ohio State University participated in this study in partial fulfillment of an introductory psychology course requirement.

Data Collection

Respondents made two visits to our laboratory, separated by approximately one month. During both visits, respondents were seated

individually in small rooms where they completed the initial or follow-up questionnaire.

Measures

Respondents reported their political party identification, ideological orientation, and attitudes on defense spending, U.S. involvement in Central America, and aid to the poor. Two questionnaire formats, partially-labeled and fully-labeled/branching, were again constructed for each attitude (for question wordings, see the Appendix). After completing their attitude reports during the second session, respondents indicated whether they felt they could remember their answers from the previous session.

Each respondent was randomly assigned to receive either partially-labeled or fully-labeled/branching questions during both visits.

Assessment of Reliability

In this study, reliability was again assessed by the percentage of respondents who provided the same answer to each question during both sessions.

Results

The Effect of Labeling/Branching

The traditional NES party identification measure was again significantly more reliable (77.8%) than the traditional NES measures of attitudes on defense spending (43.9%, $\chi^2(1)=9.44$, $p<.003$), U.S. involvement in Central America (39.0%, $\chi^2(1)=12.18$, $p<.001$), and aid to the poor (43.9%, $\chi^2(2)=9.44$, $p<.003$, see Table 7). Also as expected, verbal labeling/branching markedly increased the reliability of the attitude reports for all five items combined (63.3% vs. 43.9%, $\chi^2(1)=11.87$, $p<.001$). The largest difference in reliability occurred in

the case of party identification (33.9%, $\chi^2(1)=9.15$, $p<.003$). The defense spending items showed a somewhat smaller but nonetheless significant difference (22.8%, $\chi^2(1)=4.01$, $p<.05$). Although in the expected direction, the differences for the other two policy attitude items were not statistically significant (Central America: 16.6%, $\chi^2(1)=2.10$, $p<.15$; Aid to the Poor: 17.2%, $\chi^2(1)=2.27$, $p<.14$). As in Study 1, the reliability of ideological orientation reports was apparently unaffected by item format (-2.9%, $\chi^2(1)=0.07$, ns). Thus, the apparent effects of verbal labeling/branching observed in Studies 1 and 2 cannot be attributed completely to confusion about the which verbal labels go with which extremes on the 7-point scales.

When all attitudes were measured with the partially-labeled format, reports of party identification were just as reliable as reports of policy attitudes (see column 1 of Table 7). The reliability of party identification reports was not significantly different than the reliability of attitudes on defense spending ($\chi^2(1)=0.00$, ns), Central America attitude ($\chi^2(1)=0.20$, ns), aid to the poor ($\chi^2(1)=0.00$, ns), or the combined reliability of the three policy attitudes ($\chi^2(1)=0.03$, ns). However, when the fully-labeled/branching format was used (column 2 of Table 7), party identification reports were marginally significantly more reliable than the three policy attitude reports combined ($\chi^2(1)=3.48$, $p<.07$). However, this is completely due to the difference between the reliability of party identification reports and the reliability of Central America attitude reports ($\chi^2(1)=4.06$, $p<.05$). Party identification reports were not more reliable than defense spending attitude reports ($\chi^2(1)=1.11$, ns) or aid to the poor attitude reports ($\chi^2(1)=2.38$, ns). This again suggests that equating response

formats eliminates some but not all of the reliability difference between the traditional NES measures of party identification and policy attitudes.

Memory

To investigate the role of memory, we first computed the proportion of respondents who said that they felt they could remember their answers from the first questionnaire. A large percentage (63.9%) of respondents given the branching format reported being able to do so, whereas a significantly smaller percentage (39.0%) of those given the partially-labeled format said so ($\chi^2(1)=4.74$, $p<.03$). This raises the possibility that responses to the fully-labeled/branching questions may have been more memorable than responses to the partially-labeled questions.

However, this result turns out not to threaten the validity of our conclusions, because subjects' statements about whether or not they could recall their prior answers were not related to the over-time consistency of their attitude reports. Reliability estimates for respondents who reported they were able to remember their answers from the initial questionnaire were no higher than reliability estimates for respondents who indicated they could not remember. The difference between the two groups was not statistically significant in the cases of party identification (8.8%, $\chi^2(1)=0.63$, ns), ideological orientation (8.9%, $\chi^2(1)=0.62$, ns), defense spending (9.0%, $\chi^2(1)=0.63$, ns), Central America (-6.4%, $\chi^2(1)=0.32$, ns), aid to the poor (-11.7%, $\chi^2(1)=1.06$, ns), or the combination of the five attitudes (3.3%, $\chi^2(1)=0.40$, ns). These results suggest that reliability estimates are unaffected by a respondents ability to remember their previous answers.

This result can be interpreted in at least two ways. First, question format may affect respondents' self-reports of the ability to remember previous answers but not their actual ability to remember. That is, complete verbal labeling may lead respondents to think that they can remember their prior responses better, when in fact they cannot. A second possible interpretation is that question format does affect a respondent's ability to remember previous answers (with responses to fully-labeled/branching questions being easier to remember than responses to partially-labeled questions), but subjects made no use of their recollections in reporting their attitudes during the second session. Whichever interpretation is correct, the difference in reliability between partially-labeled and fully-labeled questions cannot be accounted for by differences in the use of recalled attitudes.

STUDY 4

In order to further explore the generality of the verbal labeling effect found in Studies 1-3, we conducted a final study using yet another data collection mode: face-to-face interviewing using show-cards. Again, given the theoretical rationale for the effects observed above, we anticipated that the labeling effect would appear here as well. In this study, we also explored whether the fully-labeled/branching questions are more reliable because of the verbal labeling *per se*, the branching *per se*, or both. In order to do so, we examined the reliabilities of partially-labeled/non-branching questions, fully-labeled/non-branching questions, and fully-labeled/branching questions.

Method

Sample

A total of 106 undergraduates at The Ohio State University participated in this study in partial fulfillment of an introductory psychology course requirement.

Data Collection

Two trained interviewers conducted face-to-face interviews with the respondents. Each respondent participated in two identical interviews separated by approximately one month. Interviewers were randomly assigned to respondents, and the same interviewer conducted the initial and follow-up interviews for each respondent.

Interviewers were not informed of the hypotheses being tested. When probed after the study, neither interviewer was able to guess the hypotheses.

Measures

All respondents answered questions about their political party identification, ideological orientation, and attitudes on defense spending, U.S. involvement in Central America, and spending for social programs. Each respondent was randomly assigned to receive one of three question formats, and each respondent received the same question form during both interviews. For one-third of the respondents, the items were measured on seven-point scales with verbal labels on only the most extreme alternatives (partially-labeled format). For another one-third of the respondents, the items were measured on seven-point scales with verbal labels on every alternative (fully-labeled format). For the final one-third, the items were measured via a series of branching questions with verbal labels on every alternative (fully-

labeled/branching format). For all three formats, respondents were handed a show-card displaying the response alternatives among which they could choose each time they were asked a question. The question wordings are shown in the Appendix.

Assessment of Reliability

In this study, reliability was again assessed as the percentage of respondents who provided the same answer to a question during both interviews.

Results

The Effect of Labeling

To our surprise, the combined reliability of the five partially-labeled items (58.9%) was not significantly different than the combined reliability of the fully-labeled items (57.8%, $\chi^2(1)=0.04$, ns), and none of the differences between reliabilities for the individual items were statistically significant. Thus, verbal labeling alone had no effect on reliability.

For the partially-labeled format, the reliability estimate for party identification (62.9%) was not significantly different from the policy attitude reliability estimates (Defense spending: 68.6%, $\chi^2(1)=0.25$, ns; Central America: 62.9%, $\chi^2(1)=0.01$, ns; social programs: 42.6%, $\chi^2(1)=2.81$, $p<.10$) or the combined policy attitude reliability (58.1%, $\chi^2(1)=0.25$, ns). However, for the fully-labeled questions, party identification (77.8%) was significantly or marginally significantly more reliable than the policy attitudes combined (52.8%, $\chi^2(1)=6.89$, $p<.01$), and attitudes on Central America (50.0%, $\chi^2(1)=5.59$, $p<.03$), social programs (50.0%, $\chi^2(1)=5.59$, $p<.03$), and defense spending (58.3%, $\chi^2(1)=3.13$, $p<.10$). Thus, party

identification was more reliable than policy attitudes using the fully-labeled/non-branching questions, though not using the partially-labeled questions.

The Effect of Branching

As the figures in Table 8 make clear, branching dramatically enhanced item reliability. The combined reliability estimate for the fully-labeled/branching format (68.0%) was significantly greater than that for the fully-labeled format (57.8%, $\chi^2(1)=3.85$, $p<.05$). Attitudes regarding social programs and party identification showed the greatest gains (21.4%, $\chi^2(1)=3.41$, $p<.07$, and 16.5%, $\chi^2(1)=4.00$, $p<.05$, respectively. Ideological orientations showed a non-significant gain (15.8%), and the other policy attitude items showed none at all. Nonetheless, it seems generally that branching is in fact responsible for the apparent effects of verbal labeling in Studies 1-3.

Within the fully-labeled/branching format, party identification was significantly more reliable (94.3%) than the policy attitudes combined (59.0%, $\chi^2(1)=14.94$, $p<.001$), defense spending (60.0%, $\chi^2(1)=11.67$, $p<.001$), U.S. involvement in Central America (45.7%, $\chi^2(1)=19.66$, $p<.001$), and social programs (71.4%, $\chi^2(1)=6.44$, $p<.025$). This result suggests that party identification benefited more from the branching format than the policy attitudes did.

DISCUSSION

Methodological Implications

These studies indicate that decomposing an attitude rating question into component questions measuring direction and intensity improves the reliability of the obtained attitude reports. This decomposition process especially improves the reliability of responses by individuals

with moderate levels of formal education and by individuals with moderately valanced attitudes. We have demonstrated this effect in telephone surveys, a self-administered survey, and a face-to-face interview, with samples of college undergraduates, a sample of residents of Columbus, Ohio, and a national sample, and for a range of political attitudes, including party identification and various policy attitudes. Therefore, the effect seems to be quite robust and prevalent across contexts.

The results of Study 4 suggest that verbal labeling on response options does not noticeably improve item reliability. However, it seems inappropriate to assume that verbal labeling definitely has no effect on reliability, given that some previous studies have demonstrated an effect (Madden, 1960; Peters & McCormick, 1966; Zaller, 1988; c.f. Finn, 1972). We look forward to seeing the results of future studies along the lines of our Study 4 in order to more conclusively resolve this issue. In the meantime, it seems clear that branching certainly does improve attitude measurement reliability and should be implemented whenever possible in surveys.

Study 1 revealed that the greatest gains in reliability due to question format occurred for respondents with moderate attitudes. Although we initially attributed this to the effect of verbal labeling, Study 4's results undermine this conclusion. An alternative interpretation focuses on the number of adjacent response alternatives respondents must choose between. In the fully-labeled format, all response options were presented to respondents at once, and they had to consider and choose among all of them. In contrast, the branching format prevented respondents from having to make choices between some

adjacent response options. Specifically, respondents were never asked directly to choose between points 2 and 3 or between points 5 and 6 on a 1-7 scale. For example, in the case of party identification, respondents were never asked directly to indicate whether they were independents leaning toward a party or were weak identifiers with that party. Making this differentiation among respondents using a sequence of branching questions did so more precisely than respondents could when asked to do so themselves in response to a single fully-labeled question. Because this sort of choice is most relevant to respondents who place themselves at points 2, 3, 5, or 6, the effect of branching should be greatest for these individuals, and this is just what we observed in Study 1.

A Note on Methodology

One aspect of the methodology of this study deserves special consideration. In all of the above discussion, we have assumed that our estimates of the over-time consistency of attitude reports reflect attitude measurement reliability. But, as we noted at the start of this paper, such figures actually reflect both measurement unreliability and the amount of true attitude change that takes place during a particular time period. Therefore, our estimates of reliability underestimate each item's true reliability if some attitude change occurred. In contrast, these figures would be biased upward if respondents remembered their initial answers to the attitude questions and simply repeated those answers during the second interview.

The one-to three-month interval between the two interviews was probably long enough to preclude respondents from vividly remembering their responses during the second interview, and the evidence from Study

3 supports this assumption. We are therefore not concerned about bias due to memory. The time interval was probably also short enough to preclude much meaningful attitude change from occurring between the two interviews, given that previous studies have documented surprisingly little attitude change occurring during much longer time periods (Alwin, Cohen, & Newcomb, 1990; Krosnick, 1988; Krosnick & Alwin, 1989a, 1989b). However, at least some small amount of attitude change probably did occur, so our estimates of reliability are probably at least slight underestimates.

Does this fact complicate interpretation of the effects of our labeling manipulations? No. If attitude change did occur, it most likely occurred in equal amounts in the various groups of respondents who received the different versions of the questions, because respondents were randomly assigned to these groups. Therefore, differences between the groups in terms of the reliability estimates are almost certainly not distorted by the presence of attitude change.

Implications for Political Science

Our findings challenge past interpretations of observed differences between party identification and policy attitudes in terms of over-time consistency in NES surveys (Converse, 1964; Converse & Markus, 1979; Markus, 1982). Previous investigators have attributed this difference to greater change in policy attitudes. However, the over-time consistency of attitude reports is a joint function of the amount of attitude change that occurs and the reliability of the measure (e.g., Alwin, 1973). And our results make clear that the branching method used to measure party identification in the NESs enhanced its reliability relative to policy attitudes. Much of the observed difference in over-

time consistency should therefore be attributed to differences in the reliability of the measures, not to stability of the underlying attitudes. Consequently, previous studies have overestimated the difference between the stabilities of party identification and policy attitudes partly because of the difference between these items' formal properties. Once all of the reliability differences are taken into account, party identification appears to be no more stable over long time periods than policy attitudes are (Krosnick & Alwin, 1989a). In fact, party identification and policy attitudes alike appear to be highly stable over long periods, though not perfectly so (see, e.g. Alwin, Cohen, & Newcomb, 1990; Krosnick & Alwin, 1989a).

Reliability affects not only the over-time consistency of attitude reports but also the amount of impact an attitude appears to have on other psychological variables in multivariate analyses. Interestingly, party identification has been shown in many analyses of NES data to be "the single most important determinant of individual voting decisions (Kinder & Sears, 1985, p. 686; see also, e.g., Campbell, Converse, Miller, & Stokes, 1960)." However, because party identification has been measured more reliably than policy attitudes in the NESs, the apparently stronger impact of party identification on variables such as vote choice must be attributed at least partly to this measurement artifact. Thus, numerous previous studies have almost certainly underestimated the impact of policy attitudes on vote choice relative to the impact of party identification.

Equating the response formats of party identification and policy attitude measures does not necessarily equate the reliabilities of these measures. We found some instances in which party identification reports

were more reliable than policy attitude reports, even when the same response format was used to measure them (see also Beck & Parker, 1985). This suggests that there are additional sources of differences in reliability between party identification and policy attitudes. As we argued above, these possible sources include the prevalence of non-attitudes and the ambiguity of internal psychological cues associated with the attitudes. Until differences between party identification and policy attitudes measures in terms of reliability are eliminated it is inappropriate to compare estimates of the over-time consistency of these attitudes or to compare the impact that these attitudes have on other variables.

Implications Regarding the 1950s NES

Our findings indicate that differences in the formal structure of attitude survey items led previous investigators to conclude that party identification is more stable than policy attitudes in the 1972-1976 and 1980 NESs. It is important to acknowledge, though, that verbal labeling and branching cannot account for the same finding in the 1956-1960 NES data. In that survey, both party identification and policy attitudes were measured with sequences of fully-verbally labeled branching questions. If branching were the only determinant of reliability, we would expect the 1950s party identification and policy attitude measures to be equally reliable. But, in fact, the 1950s party identification measure was more reliable than the 1950s policy attitude measures (Krosnick & Alwin, 1989a).

This difference in reliability may be due instead to two other differences between the formal structures of these questions: the number of response alternatives offered and the wording of those alternatives.

In the 1950s policy attitudes questions, respondents were read a statement and were asked to indicate whether they agreed strongly, agreed but not very strongly, were unsure or felt it depends, disagreed but not very strongly, or disagreed strongly. Thus, these questions offered only 5-point scales, in comparison to the 7-point party identification scales. Many studies indicate that seven-point scales are more reliable than scales with more or fewer numbers of points (Alwin & Krosnick, 1989; Andrews, 1984; Bendig, 1953; Birkett, 1986; Champney & Marshall, 1939; Finn, 1972; Symonds, 1924), so this difference may partly account for the relatively lower reliability of the 1950s policy attitude items.

These items' lower reliabilities may also be partly explained by acquiescence response bias, the tendency for some respondents to agree with any statement, regardless of its content (Lenski & Leggett, 1960; Schuman & Presser, 1981). Because the 1950s policy attitude measures involved agree/disagree response alternatives, we would expect these items to be subject to this response bias, and Jackson (1979) found evidence of just such bias in responses to these items. Relatively few respondents typically acquiesce on any given item (Lenski & Leggett, 1960; Schuman & Presser, 1981), and recent research indicates that the tendency to acquiesce is an unstable property of individuals (Hui & Triandis, 1985). That is, a person who evidences acquiescence response bias during one interview is not especially likely to evidence it during a subsequent interview a month later. It therefore seems appropriate to think of acquiescence response bias as contributing measurement error to attitude reports. Furthermore, because this error is unstable over time, it seems appropriate to think of it as increasing unreliability.

Therefore, the use of a 5-point agree/disagree response scale may account for the decreased reliability of the 1950s NES policy attitude measures relative to the measure of party identification. We look forward to future experimental research testing this speculation.

Conclusion

In conclusion, our findings suggest that researchers should be especially cautious when comparing the over-time consistencies or effects of attitudes if those attitudes were measured using items with different formal properties. Furthermore, our findings suggest that survey researchers should always use branching formats in order to maximize the reliability of their measurements. On a more substantive note, our findings add to the accumulating body of evidence (for a review, see Krosnick & Alwin, 1989a) indicating that political scientists should reconsider the widely-held assumption that party identification is more stable over time and has more impact on political cognition than any other political attitude.

Footnotes

1. It is impossible to be certain about why the response rate for this study was relatively low. It is probably resulted from the combination of our decision to make only two call-backs to a given telephone number before giving up (which presumably yielded a somewhat unusual sample being contacted) and our interviewers' lack of experience at convincing respondents to participate.
2. This test was performed using only the party identification and policy attitude measures, because the ideology question was not subject to variation in the presence or absence of a "don't know" filter. The effect of labeling on the reliability of ideology reports did not vary depending upon whether the end-points of the ideology scale were labeled "extremely" or "very".

References

Abramson, P. R. (1983)

Political Attitudes in America: Formation and Change. San Francisco:
W. H. Freeman and Company.

Alwin, D. F. (1973)

"Making inferences from attitude-behavior correlations." Sociometry
36:253-278.

Alwin, D. F., R. L. Cohen, & T. M. Newcomb (1990)

The Women of Bennington: A Study of Political Orientations Over the
Life-Span. Madison, WI: University of Wisconsin Press, forthcoming.

Alwin, D. F., and J. A. Krosnick (1989)

"The reliability of attitudinal survey data: The impact of question
and respondent characteristics." Paper to be presented at the 1989
annual meetings of the American Association for Public Opinion
Research, St. Petersburg, Florida.

Andrews, F. M. (1984)

"Construct validity and error components of survey measures: A
structural modeling approach." Public Opinion Quarterly 48:409-442.

Armstrong, J., W. Denniston, and M. Gordon (1975)

"The use of the decomposition principle in making judgments."
Organizational Behavior and Human Performance 14:257-263.

Beck, P. A., and S. Parker (1985)

"Consistency in political thinking." Political Behavior 7:37-56.

Bendig, A. W. (1953)

"The reliability of self-ratings as a function of amount of verbal anchoring and of the number of categories on the scale." *Journal of Applied Psychology* 37:38-41.

Birkett, N. J. (1986)

"Selecting the number of response categories for a Likert-type scale." *Proceedings of the American Statistical Association 1987 Annual Meetings, Section on Survey Methods.*

Bishop, G. F., R. W. Oldendick, and A. J. Tuchfarber (1980)

"Experiments in filtering political opinions." *Political Behavior* 2:339-369.

Bishop, G. F., R. W. Oldendick, and A. J. Tuchfarber (1983)

"Effects of filter questions in public opinion surveys." *Public Opinion Quarterly* 47:528-546.

Bishop, G. F., R. W. Oldendick, A. J. Tuchfarber, and S. E. Bennett (1979)

"Effects of opinion filtering and opinion floating: Evidence from a secondary analysis." *Political Methodology* 6:293-309.

Campbell, A., P. E. Converse, W. E. Miller, and D. E. Stokes (1960)

The American Voter. New York: Wiley.

Champney, H., and H. Marshall (1939)

"Optimal refinement of the rating scale." *Journal of Applied Psychology* 23:323-331.

Converse, P. E. (1964)

"The nature of belief systems in mass publics." In D. E. Apter (Ed.), *Ideology and Discontent.* Glencoe, IL: Free Press.

Converse, P. E., and G. B. Markus (1979)

"Plus ca change ...: The new CPS election study panel." *American Political Science Review* 73:32-49.

Duncan, O. D. (1975)

Introduction to Structural Equation Models. New York: Academic Press.

Fazio, R. H. (1986)

"How do attitudes guide behavior?" In R. M. Sorrentino and E. T. Higgins (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior.* New York: Guilford.

Fazio, R. H., J. Chen, E. C. McDonel, and S. J. Sherman (1982)

"Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association." *Journal of Experimental Social Psychology* 18:339-357.

Finn, R. H. (1972)

"Effects of some variations in rating scale characteristics on the means and reliabilities of ratings." *Educational and Psychological Measurement* 32:255-265.

Hui, C. H., and H. C. Triandis (1985)

"The instability of response sets." *Public Opinion Quarterly* 49:253-260.

Jackson, J. E. (1979)

"Statistical estimation of possible response bias in closed-ended issue questions." *Political Methodology* 6:393-423.

Kinder, D. R., and D. O. Sears (1985)

"Public opinion and political action." In G. Lindzey and E. Aronson (Eds.), *Handbook of Social Psychology.* New York: Random House.

Krosnick, J. A. (1988)

"Attitude importance and attitude change." *Journal of Experimental Social Psychology* 24:240-255.

————— (1989)

"Attitude importance and attitude accessibility." *Personality and Social Psychology Bulletin* 15:297-308.

————— (1990)

"Response strategies for coping with the cognitive demands of survey questions." *Applied Cognitive Psychology*, in press.

Krosnick, J. A., and D. F. Alwin (1989a)

"The stability of political preferences: Comparisons of symbolic and non-symbolic attitudes." Paper presented at the annual meetings of the American Political Science Association, Atlanta, Georgia.

————— (1989b)

"Aging and susceptibility to attitude change." *Journal of Personality and Social Psychology* 57:416-425.

Lenski, G. E., and J. C. Leggett (1960)

"Caste, class, and deference in the research interview." *American Journal of Sociology* 65:463-467.

Madden, J. M. (1960)

"A comparison of three methods of rating scale construction." USAF WADD Personnel Laboratory Technical Note, No. 60-262.

Markus, G. B. (1982)

"Political attitudes during an election year: A report on the 1980 NES panel study." *American Political Science Review* 76:538-560.

Peters, D. L., and E. J. McCormick (1966)

"Comparative reliability of numerically anchored versus job-task anchored rating scales." *Journal of Applied Psychology* 50:92-96.

Rosnow, R. L., and R. Rosenthal (1989)

"Statistical procedures and the justification of knowledge in psychological science." *American Psychologist* 44:1276-1284.

Schuman, H., and S. Presser (1981)

Questions and Answers in Attitude Surveys. New York: Academic Press.

Sherif, C. W., M. Sherif, and R. E. Nebergall (1965)

Attitude and Attitude Change. Philadelphia: W. B. Saunders.

Sherif, M., and C. I. Hovland (1953)

"Judgmental phenomena and scales of attitude measurement: Placement of items with individual choice of number of categories." *Journal of Abnormal and Social Psychology* 48:135-141.

Sherif, M., and C. I. Hovland (1961)

Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change. New Haven: Yale University Press.

Symonds, P. M. (1924)

"On the loss of reliability in ratings due to coarseness of the scale." *Journal of Experimental Psychology* 7:456-461.

Zaller, J. (1988)

"Vague minds vs. vague questions: An experimental attempt to reduce measurement error." Paper presented at the annual meetings of the American Political Science Association, Washington, D.C.

Table 1

Study 1

Percent of Respondents Providing Identical Responses
During Both Interviews

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
Party Identification	29.6%	69.4%	39.8%	p<.002
Ideological Orientation	48.1%	51.4%	3.3%	ns
Aid to Minorities	33.3%	58.3%	25.0%	p<.05
Defense Spending	51.9%	77.8%	25.9%	p<.04
Central America	29.6%	52.8%	23.2%	p<.07
Combined	38.5%	61.9%	23.4%	p<.006
N	27	36		

Table 2

Study 1

Percent of Respondents Providing Identical Responses
During Both Interviews Separately by Attitude Extremity

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
<u>Wave One Attitude Extremity</u>				
1 or 7 (Most Extreme)	62.2%	76.5%	14.3%	ns
2 or 6	28.0%	62.0%	34.0%	p<.01
3 or 5	12.8%	46.7%	33.9%	p<.005
4 (Most Moderate)	50.0%	61.5%	11.5%	ns
<u>Wave Two Attitude Extremity</u>				
1 or 7 (Most Extreme)	71.9%	72.2%	0.3%	ns
2 or 6	24.1%	54.4%	30.3%	p<.01
3 or 5	14.7%	46.7%	32.0%	p<.01
4 (Most Moderate)	42.5%	71.4%	28.9%	p<.005
N	27	36		

Table 3

Study 2

Percent of Respondents Providing Identical Responses
During Both Interviews

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
Party Identification	59.4	67.2	7.8	<.08
Ideological Orientation	42.1	51.7	9.6	<.06
Defense Spending	40.4	51.0	10.6	<.03
Central America	32.2	48.8	16.6	<.001
Gun Control	45.6	62.1	16.5	<.001
Combined	44.3	56.6	12.3	<.00001
N	249	241		

Table 4

Study 2

Percent of Respondents Providing Identical Responses
During Both Interviews By Education

Respondents	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
Low Education	37.7 (N=32)	50.3 (N=40)	12.6	<.05
Medium Education	40.9 (N=98)	58.1 (N=79)	17.2	<.001
High Education	48.6 (N=114)	57.0 (N=119)	8.4	<.01

Table 5

Study 2

Percent of Respondents Indicating "Don't Know" by Filter Condition

Attitude	"Don't Know" Filter		Difference	Significance
	Absent	Present		
Party Identification	0.7	11.9	11.2	<.0001
Defense Spending	2.3	12.7	10.4	<.0001
Central America	3.0	12.1	9.1	<.0001
Gun Control	4.0	6.1	2.1	ns
Combined	2.5	10.7	8.2	<.0001
N	310	314		

Table 6

Study 2

Percent of Respondents Providing Identical Responses
by Filter Condition

Attitude	"Don't Know" Filter		Difference	Significance
	Absent	Present		
Party Identification	61.7	64.7	3.0	ns
Defense Spending	45.1	45.9	0.8	ns
Central America	38.1	42.1	4.0	ns
Gun Control	51.3	55.8	4.5	ns
Combined	49.2	52.9	3.7	ns
N	310	314		

Table 7

Study 3

Percent of Respondents Providing Identical Responses During Both Interviews

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
Party Identification	43.9	77.8	33.9	<.003
Ideological Orientation	58.5	55.6	-2.9	ns
Defense Spending	43.9	66.7	22.8	<.05
Central America	39.0	55.6	16.6	<.15
Aid to the Poor	43.9	61.1	17.2	<.14
Combined	45.9	63.3	17.4	<.0006
N	41	36		

Table 8

Study 4

Percent of Respondents Providing Identical Responses During Both Interviews

Attitude	Question Form		Difference	Significance
	Fully Labeled	Branching		
Party Identification	77.8	94.3	16.5	<.05
Ideological Orientation	52.8	68.6	15.8	ns
Defense Spending	58.3	60.0	1.7	ns
Central America	50.0	45.7	-4.3	ns
Social Programs	50.0	71.4	21.4	<.07
Combined	57.8	68.0	10.2	<.05
N	36	35		

AppendixQuestion Wording

Study 1

Partially-Labeled Questions

Party Identification. "Some people consider themselves to be strong Republicans. Suppose these people are at one end of a seven-point scale, at point number 1. Other people consider themselves to be strong Democrats. Suppose these people are at the other end of the scale - at point number 7. Of course, other people think of themselves as somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?"

Ideology. "We hear a lot of talk these days about liberals and conservatives. Some people consider themselves to be extremely conservative. Suppose these people are at one end of a seven-point scale, at point number 1. Other people consider themselves to be extremely liberal. Suppose these people are at the other end of the scale - at point number 7. Of course, other people think of themselves as somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?"

Defense Spending. "There has been a lot of debate recently about defense spending. Some people believe that the U.S. should spend much less money for defense. Suppose these people are at one end of a seven point scale, at point number 1. Others feel that defense spending should be greatly increased. Suppose these people are at the other end of the scale - at point number 7. And, of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place your self on this scale?"

Central America. "There has been a lot of debate recently about United States involvement in the internal affairs of Central American countries. Some people think the United States should become much more involved in the internal affairs of Central American countries. Suppose these people are at one end of a seven-point scale, at point number 1. Others believe that the U.S. should become much less involved in this part of the world. Suppose these people are at the other end of the scale - at point number 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?"

Aid to Minorities. "Some people feel that the government in Washington should make every effort to improve the social and economic position of blacks and other minority groups. Suppose these people are at one end of a scale, at point number 1. Others feel that the government should not make any special effort to help minorities because they should help themselves. Suppose these people are at the other end - at point number 7. And, of course, some other people have opinions

somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?"

Coding. For all the partially-labeled questions, respondents were assigned the number between one and seven corresponding to the response alternative they selected.

Fully-Labeled Questions

Party Identification. Generally speaking, do you usually think of yourself as a Democrat, a Republican, an Independent, or what? [If Democrat:] "Would you call yourself a strong Democrat or a not very strong Democrat?" [If Republican:] "Would you call yourself a strong Republican or a not very strong Republican?" [If Independent, no preference, or other party:] "Do you think of yourself as closer to the Democratic party or the Republican party?" Strong Democrats were assigned a score of 1; not very strong Democrats were assigned a score of 2; Independents who felt closer to the Democratic party were assigned a 3; Independents and people who did not feel close to either the Democratic party or the Republican party were given a score of 4; Independents who felt closer to the Republican party were assigned a 5; not very strong Republicans were assigned a score of 6; and strong Republicans were assigned a score of 7.

Ideology. "We hear a lot of talk these days about liberals and conservatives. Generally speaking, would you consider yourself to be a conservative, a liberal, a moderate, or what?" [If conservative:] "Do you consider yourself to be extremely conservative or just conservative?" [If liberal:] "Do you consider yourself to be extremely liberal or just liberal?" [If moderate, middle of the road, or other:] "Do you think of yourself as closer to liberals or conservatives?" Extreme conservatives were given a score of 1; conservatives were given a score of 2; moderates who leaned toward conservatives or liberals were given a score of 4; moderates who leaned toward liberals were given a score of 5; liberals were given a score of 6; and extreme liberals were given a score of 7.

Defense Spending. "There has been a lot of debate recently about defense spending. Do you think the U.S. should spend less money on defense, more money on defense, or continue spending about the same amount on defense?" [If less:] "Would you say we should spend a lot less, somewhat less, or a little less?" [If more:] "Would you say we should spend a lot more, somewhat more, or a little more?" Respondents who said "a lot less" were assigned a score of 1; those who said "somewhat less" were given a score of 2; those who said "a little less" were given a score of 3; those who said "the same" were given a score of 4; those who said "a little more" were given a score of 5; those who said "somewhat more" were given a score of 6; and those who said "a lot more" were given a score of 7.

Central America. "There has been a lot of debate recently about United States involvement in the internal affairs of Central American countries. Do you think the U.S. should become more involved in the internal affairs of Central American countries, less involved in their

affairs, or continue being involved at about the same level?" [If more:] "Would you say we should become a lot more involved, somewhat more involved, or only a little more involved." [If less:] "Would you say we should become a lot less involved?" Respondents who said "a lot more" were assigned a score of 1; those who said "somewhat more" were given a score of 2; those who said "a little more" were given a score of 3; those who said "the same" were given a score of 4; those who said "a little less" were given a score of 5; those who said "somewhat less" were given a score of 6; and those who said "a lot less" were given a score of 7.

Aid to Minorities. "There has been much debate in recent years about the social and economic position of blacks and other minority groups. Do you think the government in Washington should try to improve the social and economic position of minorities, do you feel minorities should get ahead on their own, or would you endorse neither of these views? [If government should try:] "Would you say that the government should do a great deal to help minorities get ahead, should do some things to help them get ahead, or should do a little to help minorities get ahead?" [If get ahead on own:] "Would you say that minority groups are completely responsible for getting ahead on their own, somewhat responsible for getting ahead on their own, or only a little responsible for getting ahead on their own?" Respondents who said the government should do a great deal were assigned score of 1; those who said the government should do some things were given a score of 2; those who said the government should do a little were given a score of 3; those who said they endorsed neither view were given score of 4; those who said minority groups were a little responsible for getting ahead on their own were given a score of 5; those who said minorities were somewhat responsible were given a score of 6; and respondents who said minorities were completely responsible were given a score of 7.

Study 2

Partially-Labeled Questions (filters in bold)

Party Identification. Some people consider themselves to be strong Republicans. Suppose these people are at one end of a seven-point scale, at point number 1. Other people consider themselves to be strong Democrats. Suppose these people are at the other end of the scale, at point number 7. Of course other people think of themselves as somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale (remembering that 1 is a strong Republican and 7 is a strong Democrat), or haven't you thought much about this?

Ideological Orientation. Think about a ruler for measuring political views that people might hold from liberal to conservative. This ruler goes from one to seven. One means very/extremely liberal political views, and seven means very/extremely conservative political views. Just like a regular ruler, it has points in between, at 2, 3, 4, 5, or 6. Where would you place yourself on this ruler, remembering that 1 is very/extremely liberal and 7 is very/extremely conservative, or haven't you thought much about this?

Defense Spending. There has been a lot of debate recently about defense spending. Some people believe that the U.S. should spend a lot less money on defense. Suppose these people are at one end of a seven-point scale, at point number 1. Others feel that the U.S. should spend a lot more on defense. Suppose these people are at the other end of the scale -- at point number 7. And, of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale, remembering that point number 1 is a lot less spending on defense and point number 7 is a lot more spending on defense), or haven't you thought much about this?

Central America. Some people believe that the United States should become a lot less involved in the internal affairs of Central American countries. Suppose these people are at one end of a seven-point scale, at point number 1. Others believe that the U.S. should become a lot more involved in this part of the world. Suppose these people are at the other end of the scale -- at point number 7. And, of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale, (remembering that point number 1 is a lot less involved and point number 7 is a lot more involved in the internal affairs of Central American countries), or haven't you thought much about this?

Gun Control. In recent years, there has been increasing public debate about gun control laws. Some people feel the laws covering the sale of firearms should be a lot less strict. Suppose these people are at one end of a seven-point scale, at point number 1. Others feel that the laws covering the sale of firearms should be a lot more strict. Suppose these people are at the other end of the scale -- at point number 7. Of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale (remembering that point number 1 is a lot less strict gun control laws and point number 7 is a lot more strict gun control laws), or haven't you thought much about this?

Fully-Labeled/Branching Questions (filters in bold type)

Party Identification. Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what, or haven't you thought much about this? Would you call yourself a strong (Democrat/Republican) or a not very strong (Democrat/Republican)? Do you think of yourself as closer to Republicans or to Democrats?

Ideological Orientation. Generally speaking, would you consider yourself to be a liberal, a conservative, a moderate, or what, or haven't you thought much about this? Do you think of yourself as closer to liberals or conservatives? Do you consider yourself to be very (liberal/conservative) or just (liberal/conservative)?

Defense Spending. There has been a lot of debate recently about defense spending. Do you think the U.S. should spend less money on defense, more money on defense, or continue spending about the same amount on defense, or haven't you thought much about this? Would you

say the U.S. should spend a lot (less/more) or a little (less/more) on defense? Would you lean toward spending less on defense or more on defense?

Central America. Do you think the U.S. should become less involved in the internal affairs of Central American countries, more involved in their affairs, or continue being involved at about the same level, or haven't you thought much about this? Would you say we should become a lot (less/more) involved, or a little (less/more) involved? Would you lean toward the U.S. becoming less involved or more involved?

Gun Control. In recent years, there has been increasing public debate about gun control laws. Do you feel the laws covering the sale of firearms should be made less strict, more strict, or kept as they are now, or haven't you thought much about this? Do you feel these laws should be made a lot (less/more) strict or a little (less/more) strict? Would you lean toward (making) these laws less strict or more strict?

Coding. Coding was done comparably to Study 1 as described above.

Study 3

Partially-Labeled Questions

Party Identification. Some people consider themselves to be strong Republicans. Suppose these people are at one end of a seven-point scale, at point number 1. Other people consider themselves to be strong Democrats. Suppose these people are at the other end of the scale - at point number 7. Of course, other people think of themselves as somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?

Ideological Orientation. We hear a lot of talk these days about liberals and conservatives. Some people consider themselves to be extremely conservative. Suppose these people are at one end of a seven-point scale, at point number 1. Other people consider themselves to be extremely liberal. Suppose these people are at the other end of the scale - at point number 7. Of course, other people think of themselves as somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?

Defense Spending. There has been a lot of debate recently about defense spending. Some people believe that the U.S. should spend much less money for defense. Suppose these people are at one end of a seven-point scale, at point number 1. Others feel that defense spending should be greatly increased. Suppose these people are the other end of the scale - at point number 7. And, of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?

Central America. There has been a lot of debate recently about United States involvement in the internal affairs of Central American countries. Some people think the United States should become much more involved in the internal affairs of Central American countries. Suppose

these people are at one end of a seven-point scale, at point number 1. Others believe that the U.S. should become much less involved in this part of the world. Suppose these people are at the other end of the scale - at point number 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?

Aid to the Poor. There has been some debate about government services in recent years. Some people think the government should provide many fewer services, even in areas such as health and education in order to reduce spending. Suppose these people are at one end of a seven-point scale, at point number 1. Other people feel it is important for the government to provide many more services, even if it means an increase in spending. Suppose these people are at the other end of the scale - at point number 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale?

Fully-Labeled/Branching Questions

Party Identification. Generally speaking, do you usually think of yourself as a Democrat, a Republican, an Independent, or what? If you said Democrat: Would you call yourself a strong Democrat or a not very strong Democrat? If you said Republican: Would you call yourself a strong Republican or a not very strong Republican? If you said Independent or other: Do you think of yourself as closer to the Democratic party, closer to the Republican party, or equally close to both parties?

Ideological Orientation. We hear a lot of talk these days about liberals and conservatives. Generally speaking, would you consider yourself to be a conservative, a liberal, a moderate, or what? If you said conservative: Do you consider yourself to be extremely conservative or just conservative? If you said liberal: Do you consider yourself to be extremely liberal or just liberal? If you said moderate or other: Do you think of yourself as closer to liberals, closer to conservatives, or equally close to both liberals and conservatives?

Defense Spending. There has been a lot of debate recently about defense spending. Do you think the U.S. should spend less money on defense, more money on defense, or continue spending about the same amount on defense? If you said spend less: Would you say we should spend a lot less or a little less? If you said spend more: Would you say we should spend a lot more or a little more? If you said spend same: If defense spending were to be changed, would you favor spending more on defense, less on defense, or would you oppose any change in defense spending?

Central America. There has been a lot of debate recently about United States involvement in the internal affairs of Central American countries. Do you think the U.S. should become more involved in the internal affairs of Central American countries, less involved in their

affairs, or continue being involved at about the same level? If you said more involved: Would you say we should become a lot more involved or only a little more involved? If you said less involved: Would you say we should become a lot less involved or only a little less involved? If you said same: If U.S. involvement in Central America were to be changed, would you favor more involvement, less involvement, or would you oppose any change in involvement?

Aid to the Poor. There has been some debate about government services in recent years. Do you think the government should provide fewer services, even in areas such as health and education in order to reduce spending, should the government provide more services, even if it means an increase in spending, or should the government continue to provide the same number of services it currently provides? If you said more services: Would you say that the government should provide many more services or only a few more services? If you said fewer services: Would you say that the government should provide many fewer services or only a little fewer services? If you said same: If government services were to be changed, would you favor more services, fewer services, or would you oppose any change in government services?

Coding. Coding was done comparably to Study 1 as described above.

Study 4

Partially-Labeled and Fully-Labeled Questions

Party Identification. Some people consider themselves to be strong Democrats, Suppose these people are at one end of a seven-point scale, at point number 1. Other people consider themselves to be strong Republicans. Suppose these people are at the other end of the scale, at point number 7. Of course, other people think of themselves as somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale? (Labels: 2. Not very strong Democrat, 3. Independent but closer to Democrats, 4. Independent or equally close to Democrats and Republicans, 5. Independent but closer to Republicans, 6. Not very strong Republican).

Ideological Orientation. We hear a lot of talk these days about liberals and conservatives. Some people consider themselves to be extremely conservative. Suppose these people are at one end of a seven-point scale, at point number 1. Other people consider themselves to be extremely liberal. Suppose these people are at the other end of the scale - at point number 7. Of course, other people consider themselves as somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale? (Labels: 2. Conservative, 3. Moderate but closer to Conservatives, 4. Moderate or equally close to both Conservatives and Liberals, 5. Moderate but closer to Liberals, 6. Liberal).

Defense Spending. There has been a lot of debate recently about defense spending. Some people believe that the United States should spend much less money for defense. Suppose these people are at one end

of a seven-point scale, at point number 1. Other people believe that the United States should spend much more money on defense. Suppose these people are the other end of the scale, at point number 7. And, of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale? (Labels: 2. A little less, 3. Lean toward spending less, 4. Oppose any change in spending, 5. Lean toward spending more, 6. A little more).

Central America. There has been a lot of debate recently about United States involvement in the internal affairs of Central American countries. Some people believe that the United States should become much less involved in the internal affairs of Central American countries. Suppose these people are at one end of a seven-point scale, at point number 1. Other people believe that the United States should become much more involved in the internal affairs of Central American countries. Suppose these people are at the other end of the scale, at point number 7. And, of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale? (Labels: 2. A little less, 3. Lean toward becoming less involved, 4. Oppose any change in involvement, 5. Lean toward becoming more involved, 6. A little more).

Social Programs. There has been some debate about government services in recent years. Some people think the government should provide many fewer services, even in areas such as health and education in order to reduce spending. Suppose these people are at one end of a seven-point scale, at point number 1. Other people think the government should provide many more services, even if it means an increase in spending. Suppose these people are at the other end of the scale, at point number 7. And, of course, other people have opinions somewhere in between, at points 2, 3, 4, 5, and 6. Where would you place yourself on this scale? (Labels: 2. A little fewer services, 3. Lean toward providing fewer services, 4. Oppose any change in the number of services provided, 5. Lean toward providing more services, 6. A few more services).

Fully-Labeled/Branching Questions

Party Identification. Generally speaking, do you usually think of yourself as a Democrat, a Republican, an Independent, or what? [If Democrat] Would you call yourself a strong Democrat or a not very strong Democrat? [If Republican] Would you call yourself a strong Republican or a not very strong Republican? [If Independent or other] Do you think of yourself as closer to Democrats, closer to Republicans, or equally close to both Democrats and Republicans?

Ideological Orientation. We hear a lot of talk these days about liberals and conservatives. Generally speaking, would you consider yourself to be a conservative, a liberal, a moderate, or what? [If conservative] Do you consider yourself to be extremely conservative or just conservative? [If liberal] Do you consider yourself to be extremely liberal or just liberal? [If moderate or other] Do you think

of yourself as closer to conservatives, closer to liberals, or equally close to both conservatives and liberals?

Defense Spending. There has been a lot of debate recently about defense spending. Do you think the U.S. should spend less money on defense, more money on defense, or continue spending about the same amount on defense? [If less] Would you say we should spend a lot less or a little less? [If more] Would you say we should spend a lot more or a little more? [If same] Would you lean toward spending less on defense, more on defense, or would you oppose any change in defense spending?

Central America. There has been a lot of debate recently about United States involvement in the internal affairs of Central American countries. Do you think the U.S. should become less involved in the internal affairs of Central American countries, more involved in their affairs, or continue being involved at about the same level? [If less] Would you say we should become a lot less involved or only a little less involved? [If more] Would you say we should become a lot more involved or only a little more involved? [If same] Would you lean toward the U.S. becoming less involved, more involved, or would you oppose any change in involvement?

Social Programs. Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people think it is important for the government to provide more services, even if it means an increase in spending. Do you think the government should provide fewer services, provide more services, or should the government continue to provide the same number of services it currently provides? [If more] Do you think the government should provide many more services or a few more services? [If fewer] Do you think the government should provide many fewer services or only a little fewer services? [If same] Do you lean toward the government providing fewer services, more services, or do you oppose any change in the number of services the government provides?

Coding. Coding was done comparably to Study 1 as described above.

MEMO

To: NES Planning Committee and Board

From: Jon Krosnick and Matt Berent

Date: May 30, 1990

Re: Additional Analyses of the 1990 Pilot Data

At its last meeting, the Board asked that we do some additional analyses of the 1989 NES pilot data to attempt to further rule out plausible alternative explanations for our evidence indicating that branching and verbal labeling increase the reliability of attitude reports.

We have been able to do three additional types of analyses. First, a concern was raised about whether the measure of reliability that we used was well-suited to the task. This method involved calculating the percent of respondents who gave identical answers to the same question at both waves of the survey. We have now estimated the reliability of our measures using a variety of other methods.

Shown at the top of Table 2 are the old results, and at the bottom are the new ones. No matter how one calculates reliability, whether it be Gamma or Tau-b or Phi or Pearson Product-Moment Correlations or unstandardized regression coefficients, it is apparent that the branching/labeled format produced more consistent responses than the non-branched/partially-labeled format.

The second additional analyses we conducted attempted to further rule out the memory explanation for our results. This explanation posits that, during wave two, respondents were better able to recall their prior answers to the branching questions and simply repeated those prior answers, thus artificially inflating apparent reliability. As it turns out, four of our attitudes (party identification, ideological orientation, defense spending, and involvement in Central America) were measured in the 1988 NES ^{pre}post-election survey, so we could estimate the consistency of responses in the pilot study with answers given in these prior surveys months earlier.

Table 1 displays the results of these analyses. Bear in mind here that all respondents were asked the same question in 1988, but different respondents were asked different question in 1989 (either branching/labeled or non-branching/partially-labeled). In 1988, the party identification question was a fully-labeled branching question; the ideological orientation question was a fully-labeled non-branching question, the defense spending question was a partially-labeled non-branching 7-point scale, and the Central America question was a 3-point full-labeled, non-branching question.

The results here are again consistent with expectations. There is greater consistency between the 1988 questions and the branching/labeled 1989 questions than between the 1988 questions and the non-branching/partially-labeled 1989 questions.

Finally, we explore the associations between our attitude measures and general political evaluations. If, as we suspect, branching and labeling increase our measures' reliabilities, we should see stronger associations in the case of the branching/fully-labeled questions. Remarkably, it was extremely difficult to locate any general political evaluations asked comparably on all four forms of the 1989 pilot; nearly all of the questions were somehow varied across forms. The only general measure that we found was an evaluation of Reagan's performance in office, asked at the beginning of Wave One. As Table 3 shows, this item was more strongly associated with our attitudes when the latter were measured with the branching/fully-labeled questions, as compared to when the latter were measured with the non-branching/partially-labeled questions. Thus, this result is consistent with expectations.

Finally, we chose the most significant general political orientation from the 1988 post-election survey to repeat this analysis: candidate preference. Here, respondents who said they voted in the election were asked for whom they voted, and non-voting respondents were asked which candidate they preferred. As Table 4 shows, these reports of candidate preference were more strongly associated with our attitudes when measured by the branching/fully-labeled questions. Thus, we again see support for our expectations.

In sum, these additional analyses were generally reassuring.

Table 1

Study 2

Consistency of Attitude Reports between 1988 Pre-Election NES and 1989 Pilot
(Unstandardized Regression Coefficients)

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
<u>Wave One of 1989 Pilot</u>				
Party Identification	.61	.86	.25	
Ideological Orientation	.57	.60	.03	
Defense Spending	.53	.75	.22	
Central America	.92	1.07	.15	
<u>Wave Two of 1989 Pilot</u>				
Party Identification	.62	.89	.27	
Ideological Orientation	.67	.75	.08	
Defense Spending	.51	.69	.18	
Central America	.98	1.13	.15	

Table 2

Study 2

Percent of Respondents Providing Identical Responses
During Both Interviews

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
Party Identification	59.4	67.2	7.8	<.08
Ideological Orientation	42.1	51.7	9.6	<.06
Defense Spending	40.4	51.0	10.6	<.03
Central America	32.2	48.8	16.6	<.001
Gun Control	45.6	62.1	16.5	<.001
Combined	44.3	56.6	12.3	<.00001
N	249	241		

Other Measures of Reliability

Combined

Gamma	.70	.81	.11	
Tau-b	.60	.70	.10	
Phi	.41	.51	.10	All p<.01
Pearson Correlation	.68	.76	.08	
Unstd. Regression Coefficient	.67	.77	.10	

Table 3

Study 2

Predictive Validity: Association between Attitudes
and Reagan Approval Ratings (Measured in 1989 NES)
(Unstandardized Regression Coefficients)

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
<u>Wave One</u>				
Party Identification	.67	1.03		
Ideological Orientation	.55	.55		
Defense Spending	.33	.54		
Central America	.43	.42		
Gun Control	.08	.28		
<u>Wave Two</u>				
Party Identification	.72	.96		
Ideological Orientation	.31	.69		
Defense Spending	.26	.48		
Central America	.38	.52		
Gun Control	.06	.31		
<hr/>				
Combined	.35	.46	.11	p<.005
N	249	241		

Table 4

Study 2

Predictive Validity: Association between Attitudes
and Vote Choice (Voters and Non-Voters, as measured in 1988 NES)
(Unstandardized Regression Coefficients)

Attitude	Question Form		Difference	Significance
	Partially Labeled	Fully Labeled		
<u>Wave One</u>				
Party Identification	1.87	2.71		
Ideological Orientation	1.21	1.26		
Defense Spending	.93	.98		
Central America	.93	1.01		
Gun Control	.35	.70		
<u>Wave Two</u>				
Party Identification	2.12	2.67		
Ideological Orientation	1.26	1.28		
Defense Spending	.97	1.03		
Central America	.81	1.10		
Gun Control	.67	.47		
<hr/>				
Combined	.91	1.10	.19	p<.02
N	249	241		
