Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data

Colm O'Muircheartaigh
University of Chicago

Jon A. Krosnick
Ohio State University

Armin Helic
London School of Economics and Political Science

December, 2000

This manuscript was written partly while the second author was a fellow at the Center for Advanced Study in the Behavioral Sciences, supported by National Science Foundation Grant SBR-9022192. Correspondence should be addressed to Colm O'Muircheartaigh, Irving B. Harris Graduate School of Public Policy Studies, The University of Chicago, 1155 East 60th Street, Chicago, Illinois 60637 (e-mail: colm@uchicago.edu), or Jon A. Krosnick, Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, Ohio 43210 (email: krosnick@osu.edu).

Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data

AbstractAbstract

Some scholars have suggested that offering a middle alternative on a rating scale is necessary to measure opinions accurately, whereas other scholars have suggested that middle alternatives offer "easy outs" to respondents who want to avoid taking sides on an issue. In this paper, we evaluate these competing hypotheses using data from an experiment conducted in the 1992 Euro-Barometer Survey. Via structural equation modeling of responses to agree/disagree items measuring attitudes toward science and technology, we found that offering a middle alternative reduces the amount of random measurement error in the responses, thereby increasing reliability, while not affecting the validity of attitude measurements. This suggests that middle alternatives should be included in rating scales in order to maximize data quality. We also found evidence of acquiescence response bias in answers to the agree/disagree items; while unrelated to the presence of a middle alternative, this bias was stronger among older, less educated, and female respondents. And controlling for this bias greatly improved the apparent validity of attitude items.

Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data

Many survey questions measuring attitudes present respondents with a continuum along which they must place themselves. In constructing the set of response alternatives for such a question, a researcher must decide how many points to identify on the response scale and what labels to apply to these points. And in doing so, one hopes to maximize the reliability and validity of the data obtained while acquiring as much information as possible about subtle differences between different people's points of view.

Judging from research practice, there appears to have been little consensus regarding rating scale length. For example, Likert's (1932) classic attitude measurement technique uses 5-point rating scales, whereas Osgood, Suci, and Tannenbaum's (1957) semantic differential uses 7-point scales, and Thurstone's (1928) equal-appearing interval method uses 11-point scales. When American survey respondents have been asked to grade the U.S. President's job performance, it is sometimes on a 4-point scale ranging from "excellent" to "poor" (in Louis Harris surveys) and other times on a 5-point letter-grade scale from A to F (in surveys done by the Washington Post). And since the 1950s, the National Election Studies have included rating scales of many lengths, some with three points (e.g., increase, decrease, or keep the same as now), others with 4 points (e.g., strongly disapprove, disapprove, approve, and strongly approve), still others with 5 points (e.g., extremely important, very important, somewhat important, not very important, not

important at all), and others being 101-point feeling thermometers tapping attitudes toward political candidates and social groups.

Research practice seems to be nearly unanimous in one regard, though: most of the scales listed above offer odd numbers of points, meaning that they allow respondents to place themselves at a midpoint to indicate neutrality, indifference, or ambivalence. Yet surprisingly, this popular practice is in conflict with the advice of many questionnaire design experts, who have recommended omission of midpoints (e.g., Converse & Presser, 1986; Moser & Kalton, 1972, p. 344). Are practicing survey researchers simply unaware of these experts' recommendations? Or does common practice reflect a wisdom that goes beyond the thinking of these experts?

In this paper, we reconsider this issue and report new evidence on the empirical merits of offering versus omitting midpoints on attitude rating scales. Taking advantage of an experiment conducted within the 1992 Euro-Barometer Survey, we assess whether data quality is improved or compromised by offering a rating scale midpoint. Because the items involved in this experiment involved agree/disagree rating scales, we were also able to explore the presence and impact of acquiescence response bias. We begin below with a review of relevant theory and past empirical findings, and then we turn to our data.

Background

In principle, the more points a rating scale offers, the more information a researcher can learn about distinctions between people. However, the magnitude of incremental gain decreases once a sufficient number of categories are used. For example, Aigner, Goldberger, and Kalton (1975) described statistical simulations showing that the

relation between one variable and a categorical representation of another becomes increasingly strong as the number of categories increases from 2 to 10. However, whereas the information gains were quite pronounced from 2 points up to 5 points, the gains were relatively modest between 5 and 10 points. Similar results were reported at about the same time from many other simulations (Green & Rao, 1970; Jenkins & Taber, 1977; Lehmann & Hulbert, 1972; Lissitz & Green, 1975; Martin, 1973; Ramsay, 1973).

Even if more information is gained in principle with longer rating scales, the meanings of the scale points apparently become increasingly ambiguous beyond a critical point, compromising data quality. For example, various studies have shown that cross-sectional reliability, test-retest reliability, and criterion validity increase as attitude rating scales grow from 2 points to 5 points, hold steady to about 8 points, and decline as scales grow longer than that (Birkett, 1986; Masters, 1974; Matell & Jacoby, 1971; Rosenstone, Hansen, & Kinder, 1986; Smith, 1994b). Likewise, agreement between people rating the same set of objects is greatest, and susceptibility to context effects is lowest, when rating scales have 5-8 points (Finn, 1972; Schwarz & Wyer, 1985; Wedell & Parducci, 1988; Wedell, Parducci, & Lane, 1990). Thus, it appears that scales of approximately this length are optimal for obtaining high quality data.

It is less clear from the existing literature, however, whether it is wisest to employ rating scales within this range with an odd or even number of points. This is so because most previous studies have only compared scales with odd numbers of points to one another (e.g., Finn, 1972; Rosenstone, Hansen, & Kinder, 1986; Schwarz & Wyer, 1985; Smith, 1994a; Wedell & Parducci, 1988). This practice in experimental studies suggests

that researchers may have taken for granted the value of offering scale midpoints to respondents, but such an assumption was never stated explicitly in the reports of these studies.

Furthermore, there are theoretical grounds for expecting that offering midpoints may do more harm than good to data quality. According to the theory of survey satisficing (Krosnick, 1991), the cognitive processes entailed in answering a survey question carefully and precisely (called optimizing) are quite substantial (see also Tourangeau & Rasinski, 1988). Although some respondents may find this thinking relatively easy to do and may be highly motivated accordingly, most survey respondents probably find the tasks more difficult and are only minimally motivated. According to satisficing theory, these respondents may sometimes cope by giving answers that can be easily selected without much thought and can be easily justified if necessary.

In rating scale questions, the midpoint constitutes just such an option. Rather than having to think about the merits of an issue and take sides one way or the other, respondents can simply indicate that they are neutral. So if a midpoint is offered explicitly in a question, satisficing respondents may gravitate toward it. But if the midpoint were to be omitted, respondents might be encouraged instead to think about the issue and generate a substantive opinion. Consequently, omitting the midpoint may lead to the collection of meaningful opinion reports from people if no other satisficing response option is readily apparent.

Compelling as this notion is in theory, the evidence to date testing it is equivocal at best. Many experiments document that people often select middle alternatives when

they are explicitly offered, whereas few if any volunteer that response when it is omitted (e.g., Schuman & Presser, 1981). If selection of an offered middle alternative is indeed a satisficing response strategy, then this behavior should occur most under the conditions specified by satisficing theory (i.e., among respondents low in motivation and ability and when the cognitive task of responding is difficult; Krosnick, 1991). Instead, Kalton, Roberts, and Holt (1980) found that attraction to a middle alternative was unrelated to respondent educational attainment (an indirect indicator of cognitive skills), and Narayan and Krosnick (1996) found that offering a middle alternative had the most impact on people highest in formal education. Krosnick, Narayan, and Smith (1997) found the same pattern when directly assessing cognitive skills. This limited evidence challenges the satisficing perspective and leaves open the possibility that middle alternative selection may not in fact be primarily a satisficing strategy.

In this paper, we test these hypotheses anew, by assessing the impact of offering a middle alternative on data quality as gauged in a number of ways. First, we consider the reliability of responses, assessing whether omitting the middle alternative increases the amount of random error variance in reports of attitudes. Then, we consider validity, seeing whether omitting the middle alternative compromises the degree to which individual items reflect underlying attitudes and predictors of those attitudes.

The items used in this study measured attitudes by asking respondents whether they agreed or disagreed with a series of statements, which affords an opportunity to test another prediction that follows from satisficing theory. Krosnick (1991) proposed that respondents inclined toward weak satisficing may be especially likely to manifest

acquiescence response bias, the tendency to agree with any assertion regardless of its content. If this is so, then there may be a hydraulic relationship between selection of a middle alternative and saying "agree" when answering agree/disagree questions. If a middle alternative is offered, satisficers may be attracted to it. But if the middle alternative is omitted, satisficers may instead be inclined toward the "agree" end of the response scale. Thus, there may be more acquiescence response bias in answers given when the middle alternative is omitted, thus introducing systematic measurement error. If this is true, it would be another reason to offer middle alternatives to respondents. Our design allowed us to test for this possibility.

Data

The data we analyzed come from the Euro-Barometer 38.1 (Karlheinz Reif and Anna Melich, principal investigators), a survey conducted during November, 1992. The survey involved face-to-face interviews with representative samples of people ages 15 and older living in each of the 12 member states of the European Community; separate subsamples were selected in East and West Germany and in Great Britain and Northern Ireland, yielding 14 samples in all.[1] The numbers of people interviewed in each country are as follows: Belgium: 1043, Denmark: 1000, East Germany: 1014, West Germany: 1018, Greece: 1003, Spain: 1021, France: 1008, Ireland: 1000, Italy: 1021, Luxemburg: 500, the Netherlands: 1022, Portugal: 1000, Great Britain: 1066, and Northern Ireland: 308, constituting some 13,000 interviews in all. Interviews were conducted in respondents' homes in their national languages.

Midway through the interviews, respondents were asked a series of questions

tapping their general attitudes toward science and technology. Each item offered respondents a statement and asked whether they agreed or disagreed with it. Approximately half the respondents (chosen randomly, N=6511) were asked to select an answer from among the following five alternatives, displayed to them on a show card: strongly agree, agree to some extent, neither agree nor disagree, disagree to some extent, strongly disagree. In the analyses to follow, these responses were coded 5, 4, 3, 2, and 1, respectively. The other half of the respondents (N=6513) were asked to select an answer from a set of four, omitting the "neither agree nor disagree" option. Answers were coded 5, 3.66, 2.33, and 1, respectively.

For both groups of respondents, the items were presented in two sets, the first one preceded by the following lead-in: "I would like to read you now some statements that people have made about science, technology, or the environment. For each statement, please tell me how much you agree or disagree." The wordings of 10 stems in this set were as follows:

Science and technology are making our lives healthier, easier and more comfortable.

We depend too much on science and not enough on faith.

Scientific and technological research cannot play an important role in protecting the environment and repairing it.

Scientists should be allowed to do research that causes pain and injury to animals like dogs and chimpanzees if it can produce new information about serious human health problems.

Technological progress will make possible higher levels of consumption and, at the same time, an unpolluted environment.

Because of their knowledge, scientific researchers have a power that makes them dangerous.

The application of science and new technology will make work more interesting.

Most scientists want to work on things that will make life better for the average person.

Science makes our way of life change too fast.

Thanks to science and technology, there will be more opportunities for the future generations.

Interviewers did not offer a "don't know" response option explicitly to respondents but did accept and record that answer if people gave it to any particular item.

After answering these questions, respondents were asked eleven questions on their understanding of issues related to global warming. Then, interviewers said: "Now, I would like to read you some other statements. For each statement, please tell me how much you agree or disagree." The following stems were presented:

New technology does not depend on basic scientific research.

Even if it brings no immediate benefits, scientific research which advances the frontiers of knowledge is necessary and should be supported by the government.

Scientific and technological research do not play an important role in industrial development.

New inventions will always be found to counteract any harmful consequences of scientific and technological development.

Scientific research does not make industrial products cheaper.

Only by applying the most modern technology can our economy become more competitive.

Computers have made the use of bank services more complicated.

Scientific and technological progress will help to cure illnesses such as AIDS, cancer, …

The benefits of science are greater than any harmful effects it may have.[2]

Results

Shifts in Response Distributions

Table 1 displays the proportions of people who gave particular answers to each question. The first five columns pertain to people who were offered a middle alternative, the first column reporting the percent of people who agreed or disagreed strongly with each item. The second column reports the percent of people who agreed or disagreed to some extent. The third column reports the percent of people who volunteered a "don't know" response. The fourth column reports the percent of people who said they neither agreed nor disagreed. And the fifth column reports the sum of the figures in the third and fourth columns.

The final three columns pertain to people who were not offered a middle alternative. Column 6 reports the percent of them who agreed or disagreed strongly with each item. Column 7 reports the percent of people who agreed or disagreed to some

extent. Column 8 reports the percent of people who volunteered a "don't know" response. And at the bottom of each column appears the average of all figures above it.

As is apparent from these figures, offering the middle alternative attracted sizable proportions of respondents, ranging from 6.0% for Q18 to 22.7% for Q19, and averaging 14.6%. This average effect size is quite comparable to those observed in previous experimental studies systematically varying the presence or absence of middle alternatives (e.g., Schuman & Presser, 1981; Stember & Hyman, 1949-1950). What responses did these people give when the middle alternative was omitted?

If these people genuinely had no leanings one way or the other on these issues, they might have been especially inclined to volunteer "don't know" responses under these circumstances. And indeed, the proportion of people volunteering "don't know" was greater for each of the 19 issues when the middle alternative was omitted (averaging 11.4%) than when it was offered (averaging 9.4%), and this difference was statistically significant ($F(1,13022) = 32.81$, $p<.001$). However, only an average of 2.0% of respondents did so (i.e., 11.4%-9.4%), amounting to 15% of the people who selected middle alternatives when they were offered (2.0%/14.6%). Thus, most people who selected the middle alternative when it was offered apparently did not volunteer "don't know" when the middle alternative was omitted. Rather, these people most often selected from among the offered substantive response choices.

If omitting the middle alternative induced respondents who genuinely neither agreed nor disagreed with a statement to select completely randomly among the offered response alternatives, then they should have selected extreme choices just as often as

moderate ones. But if these respondents instead offered substantively meaningful responses, these responses may have tended to stay near the scale midpoint, being moderate more often than extreme. In fact, 42% of people's substantive answers (agree or disagree) were extreme on average and 58% were moderate when the middle alternative was offered, and these percentages were nearly identical and not significantly different ($F(1, 12810)=0.41$, n.s.) when the middle alternative was not offered.

Thus, it appears that omitting the middle alternative did not lead to an increase in the relative frequency of moderate responses. Instead, people who would have chosen the midpoint had it been offered were apparently disposed to select moderate and extreme options in the same ratio as people who offered non-midpoint opinions when the middle alternative was offered. This result is in line with Schuman and Presser's (1981) finding that the relative frequency of selections of non-midpoint response alternatives is usually not altered by offering or omitting the middle alternative.

If omitting the middle alternative causes respondents who would otherwise have chosen it to acquiesce instead, then the ratio of "agree" responses to "disagree" responses should be greater when the middle alternative is offered than when it is omitted. However, this turns out not to have been the case. People's substantive responses (either agree or disagree) were 64% "agree" and 36% "disagree" on average when the middle alternative was offered, as compared to 62% "agree" and 38% "disagree" when the middle alternative was omitted. This difference is statistically significant ($F(1,12810) = 72.58$, $p<.001$), though it is relatively small.

Indeed, the tendency for people who would have selected a middle alternative had

it been offered to disagree more than agree is apparent in Table 2. The average percentage of all responses that were "agree" when the middle alternative was offered was 48.6% (see column 1), which increased by 5.8% to 54.4% on average when the middle alternative was omitted (see column 2). But the average percentage of all responses that were "disagree" when the middle alternative was offered was 27.4% (see column 3), which increased by 6.8% to 34.2% on average when the middle alternative was omitted (see column 4). Thus, when the middle alternative was omitted, the primary tendency of respondents newly answering substantively was toward disagreement rather than agreement. This challenges the idea that respondents who would have selected the middle alternative if it had been offered instead acquiesced.

Distinguishing Middle Alternative Responses From "Don't Know" Responses

If people selected middle alternatives as a way of saying they had no opinion on an issue, just as they could by volunteering a "don't know" response, then the predictors of middle alternative and "don't know" responses should be quite similar. To explore this possibility, we calculated for each respondent the proportion of the items to which he or she responded "don't know" and the proportion of items to which he or she responded with the middle alternative. We then conducted ordinary least squares regressions predicting these proportions with a series of variables likely to predict opinionation in the domain of science and technology (see Krosnick & Milburn, 1990).

The predictors included: the amount of factually correct knowledge the respondent had about science, the degree to which the respondent perceived himself or herself to be knowledgeable about science, the degree to which the respondent felt he or she

understood scientific matters, his or her interest in scientific topics, exposure to news about scientific developments, and the number of television channels received in the respondent's home (presumably a more objective indicator of exposure to scientific information), as well as educational attainment, age, and gender (for the question wordings see the Appendix). The resulting unstandardized regression coefficients are shown in Table 3.

The predictors of the propensity to say "don't know" are essentially the same when the middle alternative was omitted as when it was offered (compare columns 1 and 2 of Table 3). Just as would be expected, the measures of cognitive engagement in science were negatively related to "don't know" frequency. Such responses were most common among respondents who possessed less factually correct knowledge about science, who perceived themselves to be less knowledgeable, who felt they had poorer understanding, who were less interested, and who were less exposed to information about science. Consistent with much prior research, "don't know" responses were more common among less educated respondents, older respondents, and women rather than men (see Krosnick & Milburn, 1990).

In sharp contrast, the predictors of middle alternative responses were quite different (see column 3 of Table 3). First, whereas all but one of the predictors had significant or marginally significant effects on "don't know" responses, five variables had non-significant effects on middle alternative responses (factual and perceived knowledge, exposure to news, education, and gender). Furthermore, the effects of the number of television channels received and age had reversed signs: middle alternative selections

were significantly more common among respondents who were younger and who received more television channels. Only two variables had comparable effects on both "don't know" and middle alternative responses: understanding of science and interest in news about science. Especially strikingly, the proportion of variance explained in middle alternative responses is a mere 2%, compared to 25% of the variance in "don't know" responses being explained by the same set of predictors. This clearly suggests that "don't know" and middle alternative responses are not manifestations of the same underlying cognitive state (i.e., lack of opinion).

Impact of the Middle Alternative on Reliability and Validity

If omitting the middle alternative led respondents who would have selected it to answer randomly, this would have lowered the apparent reliability of the items in tapping underlying attitudes toward science. To evaluate this possibility, we computed the reliability of the items via Cronbach's (1951) alpha, which was .64 for people who were offered the middle alternative and .65 for people not offered it.[3] These figures are not significantly different from one another ($F=1.02$, $p=.721$), and the direction of the observed difference is at odds with the notion that omitting the middle alternative induces random responding and thereby lowers reliability.

To conduct an initial assessment of the validity of the attitude measures, we focused on predictive validity. The Euro-Barometer survey included another question that measured all respondents' attitudes toward science identically. For this question, respondents were given a card listing the following professions: judges, medical doctors, lawyers, scientific researchers, businessmen, journalists, bankers, engineers, architects.

Respondents were then asked: “Which one of the following professions do you respect the most? And the second most? And which one do you respect the least?” An index of respect for scientists was created as follows: 1=respect scientists most; .66=respect scientists second most; .33=respect scientists third through eighth; 0=respect scientists the least. The stronger the association between this item and the index derived from the agree/disagree measures of attitudes, the more valid the latter presumably is.

If omitting the middle alternative induced random responding, then this association should have been stronger when the middle alternative was offered. However, the unstandardized effect of respect for scientists on the attitudes toward science index was .17 when the middle alternative was offered ($p<.001$, $N=6417$) and .21 when the middle alternative was omitted ($p<.001$, $N=6413$). These coefficients are not significantly different from one another ($z=1.25$, n.s.).

We took a second approach to assessing predictive validity as well, involving three demographic variables: age, education, and gender. Previous research has suggested that attitudes toward science and technology are more positive among younger people, among more educated people, and among men than women (e.g., Miller, 1983). We therefore expected to see these associations in the Euro-Barometer data as well, and the stronger they were, the more valid the attitude measure presumably was.

The results of regressions predicting attitudes toward science and technology with the three demographics appear in Table 4. Whereas the effect of age was equal in the two groups of respondents ($b=-.01$) and the effect of education was slightly but non-significantly larger among people not offered the middle alternative ($b=.11$ vs. .12;

z=0.44, n.s.), the effect of gender was marginally significantly larger among people offered the middle alternative than among those not offered it (b=-.13 vs. -.10, z=1.77, p<.08). This is the first hint that offering the middle alternative may increase validity, but the marginal nature of the difference limits the power of its implications.

Controlling for Acquiescence

Taken together, this evidence suggests that omitting the middle alternative did not reduce the reliability or the validity of these items. However, all of these initial findings are likely to be distorted by acquiescence response bias. A great deal of evidence now documents that some people are especially inclined to agree with all statements offered in agree/disagree items, regardless of their content, and this bias creates small positive correlations among all items employing the same agree/disagree response format (see Krosnick & Fabrigar, 1998). Because 11 of our 19 attitude items offered statements praising science and technology and the remaining 8 items offered critical or skeptical statements, this additional source of covariation among the items would misleadingly attenuate alpha estimates of reliability and any estimates of validity. If the magnitude of acquiescence varied across the two question forms, comparisons of reliability and validity across formats would be misleading.

To test for the presence of acquiescence in these data, we used LISREL 8.14 (Joreskog & Sorbom, 1993) to conduct maximum likelihood confirmatory factor analyses of covariance matrices. We constructed these matrices using the pairwise deletion approach to handling missing data in order to maximize the number of cases contributing to the estimation of each covariance. Relatively few people failed to answer each

individual item, but many people failed to answer at least one of the 19, so listwise deletion of cases reduced the sample sizes considerably. However, when we conducted the analyses reported below with listwise-deletion matrices instead, the obtained results were very close to those we report below.

In our initial analysis, we posited one factor on which all 19 items were allowed to load, representing attitude toward science and technology, and we estimated unstandardized factor loadings separately for people who had been offered the middle alternative and those who had not. To set the metric for the latent factor, the loading of the first item, which expressed a positive attitude toward science and technology, was fixed at 1.0.

The resulting loadings, shown in the first two columns of Table 5, are not at all consistent with the notion that the latent factor represents attitude toward science and technology. As would be expected, all of the 11 statements praising science and technology loaded positively and significantly. But surprisingly, only one critical or skeptical statement loaded significantly negatively. Instead, these statements most often loaded significantly positively.

A likely reason for the surprising loadings of the negative items is acquiescence response bias, which would inflate loadings of the positive items and misleadingly weaken loadings of the negative items and could even make them positive (see, e.g., Alwin & Krosnick, 1985). We therefore estimated a second pair of models in which all items were allowed to load on the same latent factor representing attitude toward science, plus a second latent factor intended to represent acquiescence. All items were constrained

to load equally on this latter factor, an assumption required to identify the model. This is reasonable, because acquiescence is defined as a tendency to agree with any item regardless of its content, so it should account for the same amount of variance in responses to all the items. The acquiescence factor was constrained to be uncorrelated with the factor representing attitudes toward science, another assumption required in order to identify the model.

As the factor loadings in columns 3 and 4 of Table 5 illustrate, accounting for acquiescence systematically altered the items' loadings on the factor representing attitude toward science. First, all the items loaded significantly on this factor in both groups of respondents, whereas some items did not load significantly in the earlier model. Second, the items expressing negative or skeptical views of science now all had sizable, negative loadings on this factor, just as would be expected on substantive grounds. Finally, the loadings of most items expressing positive views of science were consistently weaker in this model than they had been previously, which is what one would expect, because acquiescnce would have inflated these loadings in the prior models. All this is consistent with the notion that once acquiescence is accounted for, this set of items provides an effective gauge of attitudes toward science and technology.

The loading of the items on the acquiescence factor was statistically significant both for the respondents offered the middle alternative ($b=0.33$, $p<.001$) and for those not offered the middle alternative ($b=0.32$, $p<.001$). These loadings were not significantly different from one another ($\chi^2(1)=3.03$, n.s.), suggesting that acquiescence was equally present, regardless of whether the middle alternative was offered or omitted. All this

supports the presumption that acquiescence biased responses to these items and distorted correlations among them.

Having established this, we turned to comparisons of other aspects of data quality when the middle alternative was offered and omitted. To do so, we examined the impact of constraining a series of parameters to be equal across the groups of respondents who did and did not receive the middle alternative. As the second row of Table 6 shows, constraining the variance of the latent attitude factor to be equal across the two groups did not significantly compromise the goodness-of-fit of the model ($\chi^2(1)=3.66$, n.s.). Because respondents were randomly assigned to the two groups, their true underlying attitudes should not differ, and this result is consistent with that assumption.

The test of the third row of Table 6 assessed whether the factor loadings were the same across the two groups. The χ^2 statistic of 2,554.81 (with 19 degrees of freedom) was highly statistically significant ($p<.01$), meaning that the loadings differed. An examination of the loadings in the last two columns of Table 5 indicates that 14 of the 18 items with non-fixed loadings loaded more strongly on the latent attitude factor when the middle alternative was offered than when it was omitted. This suggests that offering the middle alternative increased the validity of the items. That is, a greater amount of variance in responses to each item was attributable to the latent attitude when the middle alternative was offered than when it was omitted.

The last row of Table 6 reports a test of the equivalence of the amount of random error in responses to each item. This was accomplished by constraining each item's error variance to be the same among people offered and not offered the middle alternative. The

χ^2 of 94.78 (with 19 degrees of freedom) was highly significant ($p<.01$), indicating that there were significant differences in the error variances across groups. As shown in Table 7, 15 of the 19 error variances were larger among respondents not offered the middle alternative than among those offered it. When tested individually, 9 of the items manifested significant differences in error variance between the two groups of respondents ($p<.05$; items 2, 3, 7, 8, 9, 10, 12, 16, and 18). In all nine cases, the error variance was greater among people not offered the middle alternative than among people offered it. Therefore, omitting the middle alternative increased the amount of random measurement error in reports, compromising their reliability.

Discussion

Middle Alternatives

Using a randomized split-half experiment in a cross-national survey covering the entire European Union, we found that offering a middle alternative did not compromise data quality. Contrary to the satisficing perspective, omitting the middle alternative did not lead people to report meaningful attitudes that they would otherwise not have bothered to describe, instead taking the "shortcut" of selecting the middle alternative. Rather, omitting the middle alternative significantly decreased the validity of responses and increased the amount of random error variance in responses, suggesting that people who genuinely belonged in the middle of the scale made essentially random choices among the alternatives. We saw no evidence that omitting the middle alternative increased acquiescence, so this omission seemed not to have exacerbated systematic measurement error.

Our conclusion that middle alternative selection generally does not reflect satisficing is in harmony with a number of other findings previously reported. For example, if middle alternative selection were a form of satisficing, then selection of it should occur most often among people with more limited cognitive skills. But Narayan and Krosnick (1996) and Krosnick, Narayan, and Smith (1997) found just the opposite - more middle alternative selection among people with better cognitive skills. We found no reliable relation at all between educational attainment and middle alternative selection. These failures to find the expected negative relation of cognitive skills with middle alternative selection, coupled with our evidence that offering middle alternatives did not compromise data quality, suggests confident rejection of the claim that satisficing respondents are most inclined to grab the middle alternative.

Acquiescence

Our evidence of the presence of acquiescence response bias in answers to these agree/disagree items is consistent with much prior research. However, our demonstration is one of the first showing how significantly the apparent factor structure of a set of items can be distorted by acquiescence (see also Watson, 1992; Winkler, Kanouse, & Ware, 1982). Negatively-worded items appeared to have no validity at all in tapping attitudes toward science and technology when the possibility of acquiescence was ignored, a conclusion consistent with those of a number of prior studies of other sorts of attitudes (e.g., Berkowitz & Wolkon, 1964; Chapman & Campbell, 1957; Jackson & Lay, 1968). But when we controlled for the impact of acquiescence, the apparent validity of the negatively-worded items turned out to be just as substantial as that of the positively-

worded items. For example, in column 3 of Table 4, the average loading of the positively-worded items was .82, compared to an absolute value of .99 for the negatively-worded items. Thus, it does not appear that the negatively-worded items were inherently less valid than the positively-worded items.

However, the average amount of random error in the negatively-worded items was greater (1.45 on average in column 1 of Table 8) than in the positively-worded items (1.07 on average in column 1 of Table 8). This finding is consistent with many prior studies showing that cognitive processing of negatively-worded statements is more difficult. People take longer to answer questions involving such statements than involving positively-worded statements (e.g., Gough, 1965; Wason, 1959; Wembridge & Means, 1918), and people make more errors when answering the former questions than when answering the latter (e.g., Eifermann, 1961; Dudycha & Carpenter, 1973). Thus, our findings suggest that negatively-worded items are indeed less reliable than positively-worded items. So introducing the former items into a battery in an attempt to "balance" it also introduces greater random error. This can be corrected via techniques like the structural equation modeling approach employed here. But it is quite a bit simpler and easier to avoid agree/disagree items altogether and employ balanced forced-choice items instead, which yield data of higher quality (see Krosnick & Fabrigar, in press).

If both acquiescence and middle alternative selection are the result of satisficing, one might imagine that offering a middle alternative would reduce acquiescence. That is, people inclined to satisfice might do so either by acquiescing or by selecting the middle alternative when one is offered, and more such people would satisfice by acquiescing if a

middle alternative were not offered. But we found acquiescence to be equally present when a middle alternative was offered and when it was omitted, a finding also reported by Biggers (1977). This reinforces the conclusion that middle alternative selection is not a reflection of satisficing and suggests that acquiescence cannot be reduced by offering middle alternatives.

"Don't Know" Responses

Our evidence on the causes of "don't know" responses adds to an accumulating literature on this topic (see Krosnick & Fabrigar, in press). As have previous investigators, we found such responses to be more common among less educated, older, and female respondents (e.g., Krosnick & Milburn, 1990). Our evidence that "don't know" responses are more common among people with less knowledge about the topic, less interest in the topic, less exposure to information about the topic, and less perceived competence at understanding the topic also resonates with the findings of prior work (e.g., Falkenberry & Mason, 1978; Krosnick & Milburn, 1990; Norpoth & Buchanan, 1992; Wright & Niemi, 1983). But again, ours is among the first evidence to show that these associations are robust even when controlling for each other in a multivariate analysis.

This finding discredits the simple claim that interest leads to exposure, which in turn leads to knowledge, and higher knowledge is completely responsible for drops in "don't know" rates among people high in interest and exposure. Instead, it appears that exposure, interest, and knowledge have separable and distinct effects, a conclusion being reached by other recent research (e.g., Krosnick & Brannon, 1993). In particular, our findings suggest that the decision to say "don't know" is discouraged by (1) possessing

objectively measurable knowledge with which to concoct an answer to a question, (2) believing one has been exposed to information about the topic, (3) possessing a sense of competence at understanding the domain, and (4) possessing the desire to be opinionated about the object in question. Thus, consistent with satisficing theory, "don't know" responses do not simply reflect lack of knowledge but also seem to reflect motivation to form and perhaps report opinions (see Krosnick & Fabrigar, in press).

One practical implication of our evidence involves coding of "don't know" responses. In order to maximize the number of cases contributing to a statistical analysis, investigators have sometimes treated respondents who said "don't know" as if they had provided responses at the midpoint of a rating scale. Our evidence that middle alternative and "don't know" responses have quite different origins suggests that this coding strategy may compromise data quality rather than enhance it. That is, people who say "don't know" generally have attitudes best represented at other points along a response continuum (see Krosnick & Fabrigar, in press), so these respondents generally would not be accurately represented at the midpoint. More accurate data would be obtained by encouraging these respondents to express preferences instead of saying "don't know".

Meta-Attitudinal vs. Operative Measures

Finally, our evidence reinforces an important distinction between two categories of attitude attributes, recognized recently by Bassili (1996). When operationalizing constructs, social scientists often have the opportunity to measure the construct via self-reports tapping people's perceptions or by more direct means that do not hinge on people's perceptions. So, for example, survey interviewers can either ask people to report

how much they weigh, or the interviewers can ask respondents to step on a scale to provide a more direct measurement of weight. Likewise with regard to attitude attributes, investigators can ask people how much they know about an object (which is called a meta-attitudinal measure), or people can be asked to list all their knowledge, the volume of which the investigator can then quantify (which is called an operative measure).

Although scholars have often presumed that meta-attitudinal measures can serve as effective substitutes for operative measures, recently-acquired evidence suggests that the two can sometimes diverge quite dramatically (e.g., Kanwar, Grund, & Olson, 1990; Krosnick, Boninger, Chuang, Berent, & Carnot, 1993; Visser & Krosnick, 1998). One might imagine this means that people's perceptions are both inaccurate and irrelevant to cognitive and behavioral outcomes. But the evidence we have reported clearly contradicts this interpretation. We found that both the quantity of a person's knowledge (measured operatively) and his or her meta-attitudinal perceptions of that knowledge exerted independent influences on rates of "don't know" responses. This suggests that researchers bear in mind that both perceptions and reality can be consequential, so both may be worth measuring and incorporating in theory-building.

Conclusion

This research constitutes a next step in the quest to refine and improve survey-based measures of attitudes. By experimentally varying item format and employing statistical techniques for estimating random measurement error, systematic measurement error, and correlational validity, we have generated evidence identifying more desirable and less desirable ways to design questions. And by doing so for one of the first times with a huge sample of respondents from 12 different countries, our findings merit wide generalizability.

As work of this sort continues to accumulate, social scientists progressively improve our collective ability to acquire accurate and informative data through surveys.

And by improving the precision of our measures, we permit acquiring such data at reduced cost, because larger samples are no longer needed to generate sufficient statistical power to reliably detect substantive effects. Consequently, we look forward to seeing more such work on such widely-relevant design issues.

References

Aigner, D. J., A. S. Goldberger, and G. Kalton. 1975. "On the Explanatory Power of Dummy Variable Regressions." International Economic Review 16(2):503-510.

Alwin, D. F., and J. A. Krosnick. 1985. "The Measurement of Values in Surveys: A Comparison of Ratings and Rankings." Public Opinion Quarterly 49:535-552.

Bachman, J. G., and P. M. O'Malley. 1984. "Yea-Saying, Nay-Saying, and Going to Extremes: Black-White Differences in Response Styles." Public Opinion Quarterly 48:491-509.

Bass, B. M. 1958. "Famous Sayings Test: General Manual." Psychological Reports 4:479-497.

Bassili, J. N. 1996. "Meta-Judgmental Versus Operative Indexes of Psychological Attributes: The Case of Measures of Attitude Strength." Journal of Personality and Social Psychology, 71:637-653.

Bielby, W. T. 1986. "Arbitrary metrics in multiple-indicator models of latent variables." Sociological Methods and Research 15:3-23.

Biggers, J. L. 1977. "Use of a Neutral Point with the Dogmatism Scale." Journal of Experimental Education 45:56-58.

Birkett, N. J. 1986. "Selecting the Number of Response Categories for a Likert-Type Scale." Proceedings of the American Statistical Association 488-492.

Berkowitz, N. H., and G. H. Wolkon. 1964. "A Forced Choice Form of the F Scale-Free of Acquiescent Response Set." Sociometry 27:54-65.

Centers for Disease Control and Prevention. 1995. "Mortality Patterns – The United States, 1992." Journal of the American Medical Association 273:100.

Chapman, L. J., and D. T. Campbell. 1957. "Response Set in the F Scale." Journal of Abnormal and Social Psychology 54:129-132.

Converse, J. M., and S. Presser. 1986. Survey Questions: Handcrafting the Standardized Questionnaire. Beverly Hills, CA: Sage Publications.

Cronbach, L. J. 1951. "Coefficient Alpha and the Internal Structure of Tests." Psychometrika 16:297-334.

Dudycha, A. L., and J. B. Carpenter. 1973. "Effects of Item Format on Item Discrimination and Difficulty." Journal of Applied Psychology 58:116-121.

Eifermann, R. R. 1961. "Negation: A Linguistic Variable." Acta Psychologica 18:258-273.

Falkenberry, G. D., and R. Mason. 1978. "Characteristics of Nonopinion and No Opinion Response Groups." Public Opinion Quarterly 423:533-543.

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.

Gough, P. B. 1965. "Grammatical Transformations and Speed of Understanding." Journal of Verbal Learning and Verbal Behavior 4:107-111.

Green, P. E. and V. R. Rao. 1970. "Rating Scales and Information Recovery – How Many Scales and Response Categories to Use?" Journal of Marketing 34:33-39.

Jackson, D. N., and C. H. Lay. 1968. "Homogeneous Dimensions of Personality Scale Content." Multivariate Behavioral Research 3:321-337.

Jenkins, G. D. and T. D. Taber. 1977. "A Monte Carlo Study of Factors Affecting Three Indices of Composite Scale Reliability." Journal of Applied Psychology 62:392-398.

Jöreskog, K. G., and D. Sörbom. 1993. New Features in LISREL 8. Chicago, IL: Scientific Software International.

Kalton, G., J. Roberts, and D. Holt. 1980. "The Effects of Offering a Middle Response Option With Opinion Questions." The Statistician 29:65-79.

Kanwar, R., L. Grund, and J. C. Olson. 1990. "When Do Measures of Knowledge Measure What We Think They Are Measuring?" Advances in Consumer Research 17:603-608.

Krosnick, J. A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." Applied Cognitive Psychology 5:213-236.

Krosnick, J. A., D. S. Boninger, Y. C. Chuang, M. K. Berent, and C. G. Carnot. 1993. "Attitude Strength: One Construct or Many Related Constructs?" Journal of Personality and Social Psychology 65:1132-1151.

Krosnick, J. A., and L. A. Brannon. 1993. "The Impact of the Gulf War on the Ingredients of Presidential Evaluations: Multidimensional Effects of Political Involvement." American Political Science Review 87:963-975.

Krosnick, J. A., S. S. Narayan, and W. R. Smith. 1997. "Satisficing in Surveys: Initial Evidence." in Advances in Survey Research, eds. M. T. Braverman and J. K. Slater, pp. 29-44. San Francisco: Jossey-Bass.

Krosnick, J. A., and L. R. Fabrigar. 1998. Designing Good Questionnaires: Insights From Cognitive and Social Psychology. New York, NY: Oxford University Press.

Krosnick, J. A., and M. A. Milburn. 1990. "Psychological Determinants of Political Opinionation." Social Cognition 8:49-72.

Lehmann, D. R., and J. Hulbert. 1972. "Are Three-Point Scales Always Good Enough?" Journal of Marketing Research 9:444-446.

Lenski, G. E., and J. C. Leggett. 1960. "Caste, Class, and Deference in the Research Interview." American Journal of Sociology 65:463-467.

Likert, R. 1932. "A Technique for Measurement of Attitudes." Archives of Psychology 140:1-55.

Lissitz, R. W., and S. B. Green. 1975. "Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach." Journal of Applied Psychology 60:10-13.

Martin, W. S. 1973. "The Effects of Scaling on the Correlation Coefficient: A Test of Validity." Journal of Marketing Research 10:316-318.

Masters, J. R. 1974. "The Relationship between Number of Response Categories and Reliability of Likert-Type Questionnaires." Journal of Educational Measurement 11:49-53.

Matell, M. S. and J. Jacoby. 1971. "Is There an Optimal Number of Alternatives for Likert Scale Items? Study I: Reliability and Validity." Educational and Psychological Measurement 31:657-674.

Miller, J. D. 1983. The American People and Science Policy: The Role of Public Attitudes in the Policy Process. New York, NY: Pergamon Press.

Mirowsky, J., and C. E. Ross. 1991. "Eliminating Defense and Agreement Bias from Measures of the Sense of Control: A 2 X 2 Index." Social Psychology Quarterly 54:127-145.

Moser, C., and G. Kalton. 1972. Survey Methods in Social Investigation. London: Heinemann.

Moum, T. 1988. "Yea-Saying and Mood-of-the-Day Effects in Self-Reported Quality of Life." Social Indicators Research 20:117-139.

Narayan, S., and J. A. Krosnick. 1996. "Education Moderates Some Response Effects in Attitude Measurement." Public Opinion Quarterly 60:58-88.

Norpoth, H., and B. Buchanan. 1992. "Wanted: The Education President: Issue Trespassing by Political Candidates." Public Opinion Quarterly 56:87-99.

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. 1957. The Measurement of Meaning. Urbana, IL: University of Illinois Press.

Ramsay, J. O. 1973. "The Effect of Number of Categories in Rating Scales on Precision of Estimation of Scales Values." Psychometrika 38:513-532.

Rosenstone, S. J., J. M. Hansen, and D. R. Kinder. 1986. "Measuring Change in Personal Economic Well-Being." Public Opinion Quarterly 50:176-192.

Schuman, H., and S. Presser. 1981. Questions and Answers in Attitude Surveys. New York, NY: Academic Press.

Schwarz, N., and R. S. Wyer. 1985. "Effects of Rank Ordering Stimuli on Magnitude Ratings of These and Other Stimuli." Journal of Experimental Social Psychology 21:30-46.

Smith, T. W. 1994a. "A Comparison of Two Confidence Scales." GSS Methodological Report No. 80, National Opinion Research Center, Chicago, IL.

Smith, T. W. 1994b. "A Comparison of Two Governmental Spending Scales." GSS Methodological Report No. 81, National Opinion Research Center, Chicago, IL.

Stember, H., and H. Hyman. 1949-1950. "How Interviewer Effects Operate Through Question Form." International Journal of Opinion and Attitude 3:493-512.

Stricker, L. J., S. Messick, and D. N. Jackson. 1968. "Desirability Judgments and Self-Reports as Predictors of Social Behavior." Journal of Experimental Research in Personality 3:151-167.

Thurstone, L. L. 1928. "Attitudes Can Be Measured." American Journal of Sociology 33:529-554.

Tourangeau, R., and K. A. Rasinski. 1988. "Cognitive Processes Underlying Context Effects in Attitude Measurement." Psychological Bulletin 103:299-314.

U. S. Bureau of the Census. 1993. Current Population Reports. (Series P-20, No. 474). Washington, DC: U. S. Government Printing Office.

Visser, P. S., and J. A. Krosnick. 1998. "The Development of Attitude Strength over the Life-Cycle: Surge and Decline." Journal of Personality and Social Psychology.

Wason, P. C. 1959. "The Processing of Positive and Negative Information." Quarterly Journal of Experimental Psychology 11:92-107.

Watson, D. 1992. "Correcting for Acquiescent Response Bias in the Absence of a Balanced Scale. An Application to Class Consciousness." Sociological Methods and Research 21:52-88.

Wedell, D. H., and A. Parducci. 1988. "The Category Effect in Social Judgment: Experimental Ratings of Happiness." Journal of Personality and Social Psychology 55:341-356.

Wedell, D. H., A. Parducci, and M. Lane. 1990. "Reducing the Dependence of Clinical Judgment on the Immediate Context: Effects of Number of Categories and Type of Anchors." Journal of Personality and Social Psychology 58:319-329.

Wembridge, E. R. and E. R. Means. 1918. "Obscurities in Voting Upon Measures Due to Double-Negative." Journal of Applied Psychology 2:156-163.

Winkler, J. D., D. E. Kanouse, and J. E. Ware. 1982. "Controlling for Acquiescence Response Set in Scale Development." Journal of Applied Psychology 67:555-561.

Wright, J. R., and R. G. Niemi. 1983. "Perceptions of Issue Positions." Political Behavior 5: 209-223.

Footnotes

1. Details of the sample designs are in principle available from the Euro-Barometer staff, but are not well documented. The sample designs tend to be non-probability samples, but implementation of these designs was relatively well done.

2. Four additional items were asked in this sequence, but on face validity grounds, they did not seem to tap general positive/negative orientations toward the value of science and technology. For example, one of the stems of these items was: "Some numbers are especially lucky for some people." We therefore did not include these items in the analyses described below.

3. Items 2, 3, 6, 9, 11, 13, 15, and 17 were reverse coded before computing alpha, because their statements expressed negative or skeptical attitudes toward science and technology.

Appendix: Measures of Predictors of Middle Alternative and "Don't Know" Responses

Number of Television Channels

Q23. Do you receive television in your home, via [Show card: via your own satellite dish, via a communal satellite dish, via cable, via your own aerial, via a communal aerial]?

Coding: 3=own satellite dish, communal satellite dish, cable
2=own aerial, communal aerial
1=no television

Interest in News About Science

Q50. Let us talk about these issues in the news which interest you. For each issue I read out, please tell me if you are very interested, moderately interested, or not at all interested in it. New scientific discoveries?

Coding: 3=very interested
2=moderately interested
1=not at all interested

Perceived Knowledgeability About Science

Q51. I would like you to tell me for each of the following issues in the news if you are very well informed, moderately well informed, or poorly informed about it. New scientific discoveries?

Coding: 3=very well informed
2=moderately well informed
1=poorly informed

Exposure to News About Science

Q53a. Do you ever read articles on science in newspapers? [IF YES:] Is it regularly, occasionally, or hardly ever?

Coding: 3=regularly
2.33=occasionally
1.66=hardly ever

1=never

Q53b. Do you watch TV programs on science & technology such as [EXAMPLES FOR EACH MEMBER STATE]? [IF YES:] Is it regularly, occasionally, or hardly ever?

Coding: 3=regularly
2.33=occasionally
1.66=hardly ever
1=never

Q53c. Do you read any scientific magazines, such as [EXAMPLES FOR EACH MEMBER STATE]? [IF YES:] Is it regularly, occasionally, or hardly ever?

Coding: 3=regularly
2.33=occasionally
1.66=hardly ever
1=never

AVERAGE OF THE THREE ITEMS

<u>Amount of Factually Correct Knowledge About Science</u>

Q55. Here is a quick quiz. For each thing I say, please tell me if it is true or false. If you don't know, say so, and we will skip to the next.

The centre of the earth is very hot.
The oxygen we breath comes from plants.
Radioactive milk can be made safe by boiling it.
Electrons are smaller than atoms.
The continents on which we live have been moving their location for millions of years and will continue to move in the future.
It is the father's gene which decides whether the baby is a boy or a girl.
The earliest humans lived at the same time as the dinosaurs.
Antibiotics kill viruses as well as bacteria.
Lasers work by focusing sound waves.
All radioactivity is man-made.
Human beings, as we know them today, developed from earlier species of animals.

Q56. Does the earth go around the sun or does the sun go around the earth?

Q64. Could you please tell me where you think the hole in the ozone layer is located?

Q65. Could you please tell me if you think the following statements are true or false?

The hole in the ozone layer can cause skin cancer.
The greenhouse effect can reduce the deserts.
The greenhouse effect can raise the sea level.
Acid rain can cause damage to the forests.
Car exhausts have nothing to do with acid rain.

NUMBER OF CORRECT ANSWERS OUT OF THE POSSIBLE 18.

Understanding of Science

63. In recent years, newspapers and TV have sometimes talked about to the following issues. For each of the following, could you tell me whether you have a clear understanding of what it means, a general sense of what it means, or little understanding of what it means? Acid rain? Air pollution? Global warming? The hole in the ozone layer? The greenhouse effect?

Coding: 3=clear understanding
2=general sense
1=little understanding

AVERAGE OF ALL FIVE RATINGS.

Age

Age was measured by asking respondents how old they were, and answers were recorded in years and were then recoded to range from 0 (meaning the lowest age) to 1 (meaning the highest age).

Education

Education was measured by asking respondents how old they were when they stopped full-time education, and answers were recorded in years and then recoded to range from 0 (meaning the lowest age) to 1 (meaning the highest age).

Gender

Interviewers simply recorded the gender of respondents without asking directly. Gender was coded 1 for males and 2 for females.

Table 1: Proportions of Respondents Giving Various Answers to Questions Offering and Omitting Middle Alternatives

		Middle Alternative Offered					Middle Alternative Omitted		
Question	Topic	Strong Agree/ Disagree	Weak Agree/ Disagree	DK	Middle Alter-native	MA + DK	Strong Agree/ Disagree	Weak Agree/ Disagree	DK
Question	Topic	Disagree	Disagree	native	DK	DK	Disagree	Disagree	DK
Q1	Making life healthier, easier and more comfortable.	29.2%	54.6%	4.3%	11.9%	16.2%	32.0%	63.3%	4.7%
Q2*	Depend too much on science.	31.1	41.3	7.1	20.5	27.6	37.0	53.4	9.6
Q3*	Cannot play an important role.	39.1	40.8	8.7	11.4	20.1	46.0	43.6	10.5
Q4	Allowed to cause pain and injury to animals.	42.3	38.5	5.6	13.6	19.2	46.7	45.9	7.4
Q5	Makes possible higher levels of consumption	23.5	43.6	12.8	20.1	32.9	28.3	55.6	16.0
Q6*	Power that makes them dangerous.	32.5	44.9	8.0	14.6	22.7	37.0	53.6	9.4
Q7	Make work more interesting.	24.1	48.4	9.2	18.4	27.5	28.5	60.5	11.0
Q8	Make life better for the average person.	25.6	50.7	7.0	16.7	23.7	29.5	61.2	9.3
Q9*	Makes our way of life change too fast.	29.6	47.0	5.9	17.6	23.4	34.1	58.6	7.4
Q10	More opportunities for the future generations.	31.7	45.2	7.7	15.4	23.1	34.1	55.6	10.4
Q11*	New technology does not depend on basic scientific research.	31.7	38.2	18.4	11.8	30.1	37.0	42.2	20.8
Q12	Scientific research is necessary.	36.2	42.9	9.1	11.8	20.8	39.8	50.0	10.2
Q13*	Does not play an important role in industrial development.	44.7	37.6	9.4	8.3	17.7	50.4	39.1	10.6
Q14	New inventions counteract harmful consequences.	24.6	44.0	14.8	16.6	31.4	26.6	55.8	17.6
Q15*	Does not make industrial products cheaper.	27.8	43.3	12.8	16.1	28.9	31.9	52.3	15.8
Q16	Our economy become more competitive.	34.2	42.3	11.7	11.9	23.6	36.6	49.8	13.6
Q17*	Made bank services more complicated.	42.0	37.6	8.7	11.7	20.4	47.0	43.2	9.7
Q18	Will help cure illnesses.	51.6	36.4	6.0	6.0	12.0	53.5	39.7	6.8
Q19	Benefits greater than harmful effects.	25.2	40.4	11.6	22.7	34.3	29.4	54.5	16.0
	MEAN	33.0	43.0	9.4	14.6	24.0	37.1	51.5	11.4

*These were considered unfavorable statements about the value of science and technology, whereas the remaining statements were considered to be favorable.

Note: Numbers of cases range from 6501 to 6509 for when the middle alternative was offered and from 6496 to 6507 when the middle alternative was omitted.

Table 2: Proportions of Respondents Giving Various Answers to Questions Offering and Omitting Middle Alternatives

Question	Topic Omitted	Question	Topic	Strongly Agree/Agree Offered	Strongly Agree/Agree Omitted	Strongly Disagree/Disagree Offered Disagree	Strongly Disagree/Disagree Omitted native	DK	DK	Disagree
Q1	Making life healthier, easier and more comfortable.			74.9%	82.1%	8.8%	13.2%			
Q2*	Depend too much on science.			41.1	47.7	31.3	42.6			
Q3*	Cannot play an important role.			22.6	26.1	57.3	63.4			
Q4	Allowed to cause pain and injury to animals.			32.6	38.3	48.3	54.4			
Q5	Makes possible higher levels of consumption			31.3	37.8	35.9	46.1			
Q6*	Power that makes them dangerous.			58.8	64.9	18.6	25.6			
Q7	Make work more interesting.			56.1	64.4	16.4	24.6			
Q8	Make life better for the average person.			60.7	68.1	15.7	22.6			
Q9*	Makes our way of life change too fast.			56.7	61.8	19.9	30.8			
Q10	More opportunities for the future generations.			64.3	70.2	12.5	19.4			
Q11*	New technology does not depend on basic scientific research.			19.2	22.8	50.7	56.3			
Q12	Scientific research is necessary.			72.2	78.5	6.9	11.3			
Q13*	Does not play an important role in industrial development.			14.3	16.3	67.9	73.0			
Q14	New inventions counteract harmful consequences.			45.6	52.9	23.1	29.5			
Q15*	Does not make industrial products cheaper.			45.8	51.3	25.3	32.8			
Q16	Our economy become more competitive.			66.5	72.3	9.9	14.2			
Q17*	Made bank services more complicated.			27.1	30.8	52.5	59.4			
Q18	Will help cure illnesses.			83.6	86.5	4.4	6.7			
Q19	Benefits greater than harmful effects.			50.8	61.0	14.8	23.0			
	MEAN			48.6	54.4	27.4	34.2			

*These were considered unfavorable statements about the value of science and technology, whereas the remaining statements were considered to be favorable.

Note: Numbers of cases range from 6501 to 6509 for when the middle alternative was offered and from 6496 to 6507 when the middle alternative was omitted.

Table 3: Unstandardized Coefficient Estimates From Ordinary Least Squares Regressions Predicting Middle Alternative and "Don't Know" Responses

	Dependent Variable		
	Middle Alternative Omitted	Middle Alternative Offered	
Predictor	"Don't Know"	"Don't Know"	Middle Alternative
Amount of Factually Correct Knowledge About Science	-.34***	-.33***	-.01
Perceived Knowledgeability About Science	-.11	-.16*	.02
Understanding of Science	-.14*	-.21**	-.24***
Interest in News About Science	-.38***	-.25***	-.31***
Exposure to News About Science	-.36***	-.40***	-.03
Number of Television Channels	-.20**	-.18*	.17*
Education	-1.70*	-1.92**	-.25
Age	.81***	.93***	-1.16***
Gender	.34***	.24**	.00
R^2	.25	.25	.02
N	6141	6152	6152

*p<.05
**p<.01
***p<.001

Table 4: Unstandardized Regression Coefficients Predicting Attitudes Toward Science and Technology from Demographic Characteristics

Predictor	Middle Alternative Offered	Middle Alternative Omitted
Education	.11***	.12***
Age	-.01**	-.01***
Gender	-.13***	-.10***
R^2	.04	.03
N	6404	6403

**p<.01
***p<.001

Table 5: Loadings of Items on Latent Factor Representing Attitude Toward Science and Technology

		No Acquiescence Factor		Acquiescence Factor	
Item	Topic	Middle Alternative Offered	Middle Alternative Omitted	Middle Alternative Offered	Middle Alternative Omitted
		Topic Offered	Topic Omitted	Topic Offered	Topic Omitted
Q1	Making life healthier, easier and more comfortable.	1.00	1.00	1.00	1.00
Q2*	Depend too much on science.	-0.03	0.12*	-0.86*	-0.82*
Q3*	Cannot play an important role.	0.14*	0.20*	-1.36*	-1.07*
Q4	Allowed to cause pain and injury to animals.	0.82*	0.56*	0.53*	0.40*
Q5	Makes possible higher levels of consumption	0.90*	0.82*	0.36*	0.43*
Q6*	Power that makes them dangerous.	0.02	0.00	-0.49*	-0.43*
Q7	Make work more interesting.	1.26*	1.15*	0.99*	0.95*
Q8	Make life better for the average person.	1.12*	1.03*	0.73*	0.76*
Q9*	Makes our way of life change too fast.	-0.11*	-0.07	-0.59*	-0.53*
Q10	More opportunities for the future generations.	1.43*	1.27*	1.28*	1.13*
Q11*	New technology does not depend on basic scientific research.	0.27*	0.26*	-1.35*	-1.05*
Q12	Scientific research is necessary.	0.84*	0.90*	1.04*	1.03*
Q13*	Does not play an important role in industrial development.	0.26*	0.46*	-1.39*	-1.25*
Q14	New inventions counteract harmful consequences.	0.93*	0.87*	0.35*	0.42*
Q15*	Does not make industrial products cheaper.	0.07	0.16*	-0.65*	-0.63*
Q16	Our economy become more competitive.	1.02*	1.00*	1.00*	0.93*
Q17*	Made bank services more complicated.	0.31*	0.37*	-1.25*	-1.06*
Q18	Will help cure illnesses.	0.79*	0.81*	0.88*	0.87*
Q19	Benefits greater than harmful effects.	1.23*	1.23*	0.99*	1.10*
χ^2		4,704.03	3,994.71	2,393.53	2,086.46
df		152	152	151	151
χ^2/df		30.9	26.3	15.9	13.8
Normed Fit Index		0.58	0.60	0.78	0.79
GFI		0.89	0.91	0.95	0.96
N		5601	5383	5601	5383

*These were considered unfavorable statements about the value of science and technology, whereas the remaining statements were

considered to be favorable.

*p<.05

Table 6: Goodness of Fit Indices for Models Including Various Parameter Constraints

Model	χ^2	df	$\Delta\chi^2$	df	p
No across-group constraints	4,479.99	302			
Latent attitude factor variance equal across groups	4,483.65	303	3.66	1	n.s.
All factor loadings equal across groups	7,034.80	321	2,554.81	19	p<.01
Error variances equal across groups	4,574.77	321	94.78	19	p<.01

Note: The number of cases is 10,984 for all models.

Table 7: Error Variance Estimates From Structural Equation Models Including Acquiescence Factor

Item Omitted	Topic	Middle Alternative Offered	Middle Alternative Offered			
	Topic		Omitted	Offered	Omitted	Offered Omi
Q1	Making life healthier, easier and more comfortable.	0.69	0.69			
Q2*	Depend too much on science.	1.54	1.66			
Q3*	Cannot play an important role.	1.33	1.41			
Q4	Allowed to cause pain and injury to animals.	1.90	1.94			
Q5	Makes possible higher levels of consumption	1.39	1.41			
Q6*	Power that makes them dangerous.	1.42	1.48			
Q7	Make work more interesting.	0.98	1.05			
Q8	Make life better for the average person.	0.97	1.03			
Q9*	Makes our way of life change too fast.	1.26	1.38			
Q10	More opportunities for the future generations.	0.84	0.93			
Q11*	New technology does not depend on basic scientific research.	1.23	1.29			
Q12	Scientific research is necessary.	0.73	0.81			
Q13*	Does not play an important role in industrial development.	1.06	1.04			
Q14	New inventions counteract harmful consequences.	1.36	1.32			
Q15*	Does not make industrial products cheaper.	1.54	1.59			
Q16	Our economy become more competitive.	0.88	0.91			
Q17*	Made bank services more complicated.	1.73	1.73			
Q18	Will help cure illnesses.	0.58	0.68			
Q19	Benefits greater than harmful effects.	0.98	1.02			
	N	5601	5383			

*These were considered unfavorable statements about the value of science and technology, whereas the remaining statements were considered to be favorable.