EDUCATION MODERATES SOME RESPONSE EFFECTS IN ATTITUDE MEASUREMENT

SOWMYA NARAYAN
JON A. KROSNICK

Abstract On the basis of over 130 survey-based experiments, Schuman and Presser concluded that respondents' levels of educational attainment did not seem to be a pervasive and systematic moderator of susceptibility to response effects. Via a meta-analysis of these same data, we found that lower education was indeed associated with greater strength of seven response effects: response order effects, acquiescence, middle alternative effects not involving status quo options, no-opinion filter effects, forbid/allow effects, balance effects, and question order effects based on the norm of reciprocity. The specific patterns of relations obtained are consistent with the notion that some of these response effects may result from satisficing.

It is well established that people's reports of their attitudes are influenced by factors in addition to the attitudes themselves. Specifically, changes in the form, wording, and sequencing of questions can alter people's responses to them (Schuman and Kalton 1985; Schuman and Presser 1981; Tourangeau and Rasinski 1988). In the largest systematic study of such response effects, Schuman and Presser (1981) conducted over 130 experiments in 34 surveys of representative national and regional samples of American adults between 1971 and 1980. These experiments documented effects of question and response order, acquiescence, middle alternatives, no-opinion filters, tone of wording, question balance, and other formal aspects of questions.

In addition to documenting these effects, Schuman and Presser (1981) explored which sorts of respondents were most susceptible to

SOWMYA NARAYAN is a research associate at Gallup Marketing and Business Associates, Bombay, India. JON A. KROSNICK is associate professor of psychology and political science at Ohio State University. The authors would like to thank John Cacioppo, Richard Petty, Stanley Presser, Norbert Schwarz, and Roger Tourangeau for helpful comments on an earlier draft of this article. Correspondence should be directed to Jon Krosnick, Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, OH 43210; E-mail krosnick@osu.edu.

Public Opinion Quarterly Volume 60:58–88 0033-362X/96/6001-0006$02.50

them. In this regard, Schuman and Presser (1981) devoted special attention to testing one primary hypothesis: that more educated respondents might be less likely to manifest all varieties of response effects. Schuman and Presser suspected that such individuals might "more easily grasp the general point of a question and be less easily affected by emotionally colored words" (p. 6), find arguments offered by a question to be less novel and therefore less persuasive, and "be less deferential toward the interviewer and the questionnaire" (p. 7). For all these reasons, Schuman and Presser (1981) suspected, people with relatively little formal education might be most influenced by the inclusion or omission of particular response options, the particular words used in a question, and its response format. At the same time, however, Schuman and Presser (1981) noted that better-educated individuals might "be more sensitive to precise verbal distinctions where that is appropriate" (p. 6). Thus, these investigators thought it conceivable that education might not always moderate the magnitude of response effects or that better-educated respondents might sometimes show stronger effects.

In order to assess the relation of education to the magnitudes of response effects, Schuman and Presser (1981) estimated the impact of their many question manipulations separately among respondents with low, moderate, and high levels of education. Some of the statistical tests of the relation between education and response effect size were statistically significant and revealed the expected pattern. But the majority of such experiments failed to support their expectations. Consequently, Schuman and Presser (1981) concluded, with regard to response order effects, for example, that "there are some trends in the data that support this hypothesis, but none of the three-way interactions involving education (i.e., response by response order by education) is significant at the .05 level, and there are enough contradictory trends to render the conclusion doubtful. At this point, we do not see any consistent evidence that response-order effects vary systematically by educational level" (p. 71).

This sort of statement is typical of the conclusions these investigators reached concerning the effect of education on the likelihood of middle alternative effects ("there is little evidence for this hypothesis" [p. 171]), no-opinion effects ("most of our experiments fail to confirm the . . . hypothesis" [p. 140]), balance effects ("low education cannot be considered a factor that invariably creates susceptibility to influence [p. 195]), tone of wording effects ("there is not much sign of an interaction with education" [p. 282]), and question order effects (see pp. 23–56). There were a few exceptions to this general pattern: acquiescence effects, forbid/allow tone of wording effects, question order effects due to the norm of reciprocity, and no-opinion effects for obscure issues did seem to be reliably stronger among less educated respondents. But in general, Schuman and Presser (1981) concluded, education was not a "pervasive and systematic" determinant of whether respondents were highly susceptible to response effects (see p. 303).

We began the investigation reported below because we thought there might be reason to reconsider this conclusion. First, and most important, we found the theoretical rationale for expecting education effects to be quite compelling. In addition to Schuman and Presser's (1981) speculations, a number of plausible psychological and sociological accounts of various response effects implicate education as a likely moderator (see, e.g., Krosnick 1991). Furthermore, although Schuman and Presser (1981) reported only a few of the tests of the education relation they conducted for such effects, the patterns they did report were often in the expected direction, albeit often not statistically significant. This led us to wonder whether the item-by-item analysis approach employed by Schuman and Presser (1981) might have hampered their ability to detect a significant overall trend across all of their experiments.

In fact, the advent of meta-analysis in recent years has alerted researchers across the social sciences to the value of analyzing sets of empirical studies in tandem, rather than considering each one individually (see, e.g., Oliver 1987; Rosenthal 1988). Not only does a coordinated simultaneous analysis maximize statistical power, but it allows one to see past the specific and idiosyncratic dynamics of any particular experiment. Thus, one can evaluate, in a precise and objective manner, whether an expected relation holds across a number of independent studies by combining the effect sizes or significance levels obtained in them. The advantage here is that one can detect real effects that may have escaped detection using more traditional methods (Rosenthal 1988). Thus, the effect of education in any single experiment by Schuman and Presser (1981) might not have appeared to be reliable, but pooled results might indicate otherwise. For this reason, we conducted a meta-analysis of Schuman and Presser's (1981) experiments to reexamine the role of education in moderating response effects.

Data

DESIGN

Most of the more than 130 experiments conducted by Schuman and Presser (1981) were done in telephone surveys, and the rest were conducted via face-to-face interviews in respondents' homes. Although some were more complicated, most of the experiments involved essentially the same design: after answering an array of prior questions, half the respondents (selected randomly) were asked one version of a target question, and the other half were asked a different version of the question (for details on the study designs, see Schuman and Presser 1981).[1]

MEASURES

Response order. To assess response order effects, the order in which response alternatives were presented to respondents was varied. For example, half the respondents in one experiment received the following question: "Some people say that we will still have plenty of oil 25 years from now. Others say that at the rate we are using our oil, it will all be used up in about 15 years. Which of these ideas would you guess is most nearly right?" The other half of the respondents received the same question with the order of the alternatives reversed: "Some people say that at the rate we are using our oil, it will all be used up in about 15 years. Others say we will still have plenty of oil 25 years from now. Which of these ideas would you guess is most nearly right?" Comparison of responses to the two question forms allows one to estimate the impact of response order. Schuman and Presser (1981) conducted 17 such experiments, addressing topics such as oil companies, oil supply, divorce, aid to Vietnam, trade unions, and housing.

Acquiescence. To assess acquiescence, some respondents were asked to agree or disagree with a statement, whereas others were asked to choose between two statements expressing two mutually exclusive points of view. For instance, some respondents in one experiment were asked, "Please tell me whether you agree or disagree with this statement: 'Individuals are more to blame than social conditions for crime and lawlessness in this country.'" Other respondents were asked instead, "Which in your opinion is more to blame for crime and lawlessness in this country—individuals or social conditions?" Comparison of answers to these two question forms allows for estimation of respondents' tendencies to agree with any assertion, regardless of its content. Fourteen of Schuman and Presser's (1981) experiments were of this type, addressing topics such as women in politics, crime, and foreign policy.[2]

1. Although we did not analyze Schuman and Presser's (1981) experiments involving comparisons of open-ended to closed-ended questions (because there was no way to calibrate effect sizes for those experiments), we did analyze almost all of their other experiments. Some of the ones we analyzed were not reported in Schuman and Presser's (1981) book, but they were suitable for inclusion and therefore were incorporated. Some other experiments that were reported in the book were not included in our analyses, for reasons we shall explain below.
2. In exploring acquiescence, we did not analyze all of the possible comparisons between agree/disagree question forms and forced choice forms. To control for response order

Middle alternative effects. The tendency to select a middle alternative was investigated in experiments that included and omitted such alternatives. For example, half the respondents in one experiment were asked: "In your opinion, should the penalties for using marijuana be more strict or less strict than they are now?" The other half were asked instead: "In your opinion, should the penalties for using marijuana be more strict, less strict, or about the same as they are now?" Of interest here is the increase in the proportion of respondents endorsing the middle alternative as a result of the cue in the question legitimating that response. Schuman and Presser (1981) conducted 15 such experiments, in all of which dichotomous bipolar items were expanded to be three-category items. These items involved topics such as divorce, local education, respondents' political leanings, aid to Vietnam, and marijuana legalization.

No-opinion effects. The fourth variation that we examined involved the inclusion or omission of no-opinion response options in questions about either familiar or obscure issues. For the familiar issues, two sorts of filtering processes were examined. In the quasi-filter experiments, some respondents were asked questions like the following: "Do you feel that almost all of the people running government are smart people, or do you think that quite a few of them don't seem to know what they're doing?" The remaining respondents in this experiment were asked, "Do you feel that almost all of the people running government are smart people, or do you think that quite a few of them don't seem to know what they're doing, or do you not have an opinion on this?" Although some people volunteered that they had no opinion when asked the first question, many more selected that response when it was offered explicitly in the second. The focus here, then, is on the increase in the proportion of respondents saying they had no opinion because of the cue in the question legitimating that response. Nine of Schuman and Presser's (1981) experiments were of this sort, addressing communism, the courts, government leaders, and respondents' political leanings.

In the full filter experiments, some respondents were again simply asked for their opinions on a topic without mention of a no-opinion option. Other respondents were initially asked simply whether or not they had an opinion on an issue. Only if they said they did were they

effects, we only examined cases in which the order of the substantive response alternatives was the same on both the agree/disagree and forced choice versions. Schuman and Presser (1981, p. 228) expressed skepticism about whether some of their acquiescence experiments successfully constructed forced-choice item versions that were well-suited to comparison with agree/disagree versions of the same items. We nonetheless included some of those experiments in our analysis because no significant problems with the items were apparent to us.

asked to report those opinions by choosing among substantive response options. For example, respondents in one experiment were asked, "Here are some questions about other countries. Not everyone has opinions on these questions. If you do not have an opinion, just say so. The Arabs are trying to work for a real peace with Israel. Do you have an opinion on that? [If yes] Do you agree or disagree?" Again, the impact of this sort of filter on the frequency of no-opinion was assessed. Eleven of Schuman and Presser's (1981) experiments were of this sort, addressing Russian leaders, the Portuguese government, government power, and the Middle East.

Two of Schuman and Presser's (1981) no-opinion filter experiments involved obscure issues, with which most respondents were unlikely to be familiar. These experiments both entailed quasi-filters.

Forbid/allow. Most of the experiments under the "tone of wording" heading varied the phrasing of attitude question stems, comparing the word "forbid" to the word "allow." For example, in one experiment, some respondents were asked, "Do you think the United States should forbid public speeches against democracy?" Other respondents were asked instead, "Do you think the United States should allow public speeches against democracy?" Comparisons of responses permit one to assess the impact of such wording shifts, because if the wording shift has no effect, the proportions saying "yes—forbid" and "no—do not allow" should be equivalent. Six forbid/allow experiments were conducted by Schuman and Presser (1981), addressing pornographic movies, cigarette commercials, and speeches in favor of democracy and communism.[3]

Balance. The majority of Schuman and Presser's (1981) balance experiments compared three types of questions. Unbalanced questions presented only one side of an issue and asked for a yes/no response. For example, in one experiment, some respondents were asked, "Would you favor a law which would require a person to obtain a police permit before he could buy a gun?" Formally balanced questions simply added the formal acknowledgment of the alternative viewpoint, as in: "Would you favor or oppose a law which would require a person to obtain a police permit before he could buy a gun?" Comparisons of these types of questions allows one to assess the impact of formal balancing.

In some experiments, Schuman and Presser (1981) compared unbal-

3. Three additional experiments under the "tone of wording" heading varied subtle aspects of question wording in ways other than the use of "forbid" and "allow." Two of these experiments were on the issue of abortion, and one experiment was on the issue of atheists. The variations in wording that these experiments involved cannot be subsumed under a single heading, and only one of these experiments yielded a significant form effect. So no further analysis seemed sensible.

anced or formally balanced items to ones that increased the number of words devoted to articulating one of the opposing viewpoints, thus tipping the balance of prominence in that direction. We refer to such comparisons as assessing the impact of enhanced viewpoint salience. This was sometimes done simply by stating one point of view in more detail in one question form than the other, as in the comparison between these items: "If there is a union at a particular company or business, do you think that all the workers there should be required to be union members?" and "If there is a union at a particular company or business, do you think that all the workers there should be required to be union members, or should it be left to the individual to decide whether or not he wants to be in the union?" In other experiments, unbalanced or formally balanced items were compared to items that presented one viewpoint endorsed by an argument in its favor, as in the following sample: "Would you favor a law which would require a person to obtain a police permit before he could buy a gun, or do you think such a law would interfere too much with the rights of citizens to own guns?" Comparisons of salience-enhanced questions to unbalanced or formally balanced questions allow one to assess the impact of tipping the balance (in terms of the number of words articulating each point of view) toward one opinion or another.

Schuman and Presser (1981) conducted eight experiments comparing unbalanced items to formally balanced items, eight comparing unbalanced items to salience-enhanced items, and 16 comparing formally balanced items to salience-enhanced items. These experiments addressed topics such as fuel shortages, trade unions, abortion, gun permits, Russian leaders, Vitenam, and communism.[4]

Question order. Question order effects were assessed by manipulat-

4. Schuman and Presser (1981) treated experiments asking about sending troops if "another Vietnam" were to occur as examining tone of wording effects. However, the manipulation simply involved adding the phrase "to stop a Communist takeover" to a formally balanced question. Because this seemed to enhance the salience of one viewpoint, this manipulation seemed comparable to other experiments under the "balance" heading, so we treated these experiments as testing for balance effects. However, dropping this item from the balance analyses did not alter their conclusions at all. Schuman and Presser (1981) treated questions about removing a book written by a Communist from the public library as exploring tone of wording. However, the difference between the "Stouffer" question form and the "liberal" question form in the Communist book experiment was the addition of an argument: "Somebody else in your community says this is a free country and it should be allowed to remain." Consequently, we treated this comparison as a test of the impact of this argument, under the rubric of balance effects. Likewise, the difference between the "Stouffer" form and the "neutral" form was the addition of an argument on one side, so it too was treated under the rubric of balance effects. And for the question about an atheist (also treated by Schuman and Presser as involving tone of wording), the comparison of the "Stouffer" form and the neutral form involved the addition of an argument, so it was treated under the balance rubric as well.

ing the order in which questions were asked. For example, some respondents in one study were asked first, "Do you think the United States should let communist newspaper reporters from other countries come in here and send back to their papers the news as they see it?" and second, "Do you think a communist country like Russia should let American newspaper reporters come in and send back to America the news as they see it?" The remaining respondents were asked the same two questions in the reverse order. Any impact of this manipulation on responses reveals an order effect. Schuman and Presser (1981) conducted 23 question order experiments, addressing abortion, foreign newspaper reporters, occupational discrimination, and other topics.

Education. In all of Schuman and Presser's (1981) experiments, respondents were asked to report the highest level of education they had completed. For our analyses, respondents were divided into three groups of approximately equal size: those who had not completed high school (the low-education group), high school graduates (the medium-education group), and people who had attended at least some college (the high-education group).

Analysis

Each experiment was included in our meta-analysis only if the response effect of interest was statistically significant or marginally significant for the full sample of respondents. Not all experiments intended to yield response effects succeed in doing so, and if no effect exists in a particular study, it makes no sense to use that study to identify a moderator. Including enough such failed experiments in a meta-analysis could wipe out the expected differences across education levels because equivalent effects (i.e., of zero magnitude) would have occurred in all three education groups. Employing this criterion eliminated 10 response order experiments, three acquiescence experiments, one forbid/allow experiment, nine comparisons of unbalanced or formally balanced items to salience-enhanced items, and 10 question order experiments (for a listing of all experiments and their results, see appendix table A1).[5]

The meta-analysis was done according to a standard procedure described by Rosenthal (1988) using a computer package called *Advanced Basic Meta-Analysis* (Mullen 1989). For each experiment, we first cal-

5. In order to ascertain whether eliminating these studies affected our conclusions, we also conducted the meta-analyses including these omitted experiments. The results of this combined analysis were comparable to those reported below, though a bit weaker.

culated a χ^2 for the effect of question format, wording, or order within each education subgroup. This method involved the same log-linear approach that Schuman and Presser (1981) used in their initial analyses of the individual experiments. The resulting χ^2s were then converted to measures to effect size (Cohen's d) for the individual education groups (for an explanation of effect size metrics, see Rosenthal [1988]). A value of 0 indicates no effect of a question manipulation; a positive number indicates an effect in the same direction as the full sample; and a negative number indicates an effect in the direction opposite that of the full sample. The meta-analysis compared the average effect size for the three education groups to one another.

Because we had strong a priori expectations regarding the directions of the education effects, all significance levels reported below are one-tailed, except in the handful of cases where the observed difference ran in the direction opposite to our expectations (in which cases two-tailed p's are reported).

Because our meta-analyses were based on thousands of cases, one might be inclined to presume that they have tremendous statistical power to detect real effects. This view would suggest that one take seriously only those effects that are conventionally significant and not those that are only marginally significant. This is certainly a reasonable approach and one that we shall follow.

At the same time, however, we see merit in a more open-minded view. In conducting our meta-analyses in various different ways (e.g., slightly changing the pool of items included in an analysis), we have seen a few education effects sometimes be conventionally significant and other items be marginally significant. But we never saw effects transform from being conventionally significant to being clearly nonsignificant. Consequently, marginal significance here may well signal education effects that will turn out to be clearly reliable when meta-analyses of larger pools of items are conducted in the future. We therefore take conventionally significant effects most seriously but caution readers not to disregard marginally significant effects completely.

Results

Response order. Table 1 displays the average effect size for each type of question manipulation in each education subgroup. As expected, the average effect size for the response order experiments was greater in the low-education group (d = .36) than in the medium- (d = .15) and high-education (d = .18) groups. As the statistical tests

Table I. Effect Sizes for the Three Education Groups

Response Effect	Number of Experiments	Effect Sizes: Low-Education Group	Medium-Education Group	High-Education Group
Response order	7	.36	.15	.18
Acquiescence	11	.33	.19	.18
Middle alternative:				
All	15	.55	.59	.42
Status quo	6	.35	.46	.43
Others	9	.71	.70	.43
No opinion, familiar issues:				
Full filters	11	.60	.69	.49
Quasi-filters	9	.70	.54	.39
No opinion, obscure issues	2	.74	.47	.41
Forbid/allow	5	.45	.48	.25
Balance:				
Unbalanced versus formal balance	8	.008	.004	.007
Unbalanced and formal balance versus salience-enhanced	12	.31	.28	.22
Question order:				
All	13	.42	.39	.21
Norm of reciprocity	4	.67	.65	.15
Others	9	.30	.27	.24

NOTE.—These analyses included only those experiments with significant or marginally significant form effects for the full sample of respondents.

in the first row of table 2 show, the low-education group's effect size was significantly larger than the medium-education group's (z = 2.43, p = .01) and the high-education group's (z = 2.08, p = .01). The medium-and high-education groups did not differ (z = 0.49, p = .31).

Acquiescence. The acquiescence effect size was also larger in the low-education group (d = .33) than in the medium- (d = .19) and high-education (d = .18) groups. The low-education group's effect size was significantly larger than those of the medium-education group (z = 3.07, p < .001) and the high-education group (z = 3.21, p < .01). The

Table 2. Tests of the Significance of the Differences between Effect Sizes

Response Effect	Low versus High Education		Low versus Medium Education		Medium versus High Education	
	z	p	z	p	z	p
Response order	2.08	.02	2.43	.01	.49	.31
Acquiescence	3.21	.00	3.07	.00	.16	.43
Middle alternative:						
All	2.99	.00	.97	.16	4.51	.00
Status quo	1.50	.07	1.70	.05	.46	.32
Others	4.78	.00	.11	.46	5.10	.00
No opinion, familiar issues:						
Full filters	2.41	.01	1.76	.08	4.33	.00
Quasi filters	5.43	.00	2.73	.00	3.30	.00
No opinion, obscure issues	2.39	.00	1.88	.02	.57	.28
Forbid/allow	3.08	.00	.48	.32	3.89	.00
Balance:						
Unbalanced versus formal balance	.17	.43	.70	.24	.62	.27
Unbalanced and formal balance versus salience-enhanced	1.87	.03	.57	.29	1.36	.09
Question order:						
All	3.12	.00	.44	.33	3.21	.00
Norm of reciprocity	3.34	.00	.14	.44	3.84	.00
Others	1.01	.15	.50	.31	.61	.26

NOTE.—These analyses included only those experiments with significant or marginally significant form effects for the full sample of respondents.

medium- and high-education groups' effect sizes did not differ (z = 0.16, p = .43).

Middle alternative effects. The impact of offering middle alternatives also varied with education. The high-education group's average effect size (d = .42) was significantly smaller than those for the medium-education group (d = .59, z = 4.51, p < .01) and the low-education group (d = .55, z = 2.99, p < .01). The low- and medium-education groups were not significantly different (z = .97, p = .16).

Some of Schuman and Presser's (1981) middle alternative experiments were of a particular type, where the middle alternative endorsed maintaining the status quo in the future (e.g., keep divorce laws as they are now). Surprisingly, the average effect size of offering such a status quo alternative in the low-education group (d = .35) was smaller than the comparable effect size in the medium-education group (d = .46) and the high-education group (d = .43). However, none of these differences were statistically significant (low vs. high: z = 1.50, p = .13; low vs. medium: z = 1.70, p = .09; medium vs. high: z = 0.46, p = .65).

The remaining middle alternative experiments offered different sorts of middle alternatives (a ''middle-of-the-road'' option on an item about political ideology and an ''about right'' option on a question about whether U.S. government help to the South Vietnamese government had been adequate). The impact of offering such options was apparently greater among less educated respondents. The high-education group's average effect size (d = .43) was significantly smaller than those for the medium-education group (d = .70, z = 5.10, p < .01) and the low-education group (d = .71, z = 4.78, p < .01). The low- and medium-education groups were not significantly different (z = .11, p = .46).

No-opinion effects for familiar issues. For questions on familiar issues, the average effect of offering a full no-opinion filter was greater for the low- and medium-education groups (d = .60 and .69, respectively) than for the high-education group (d = .49). The high-education group's effect was significantly smaller than that of the low-education group (z = 2.41, p < .01) and that of the medium-education group (z = 4.33, p < .01). The medium- and low-education groups differed marginally significantly (z = 1.76, p = .08).

The quasi-filter experiments involving familiar issues revealed a somewhat different pattern. Here, the effect was again greatest among the low-education group (d = .70), smaller among the medium-education group (d = .54), and smallest among the high-education group (d = .39). The high-education group's effect was significantly smaller than that of the low-education group (z = 5.43, p < .01) and that of the medium-education group (z = 3.30, p < .01), and the medium-education group's effect was significantly smaller than that of the low-education group ($z = 2.73, p < .01$).[6]

No-opinion filter effects for unfamiliar issues. Yet another pattern appeared for the experiments involving unfamiliar issues, all of which used quasi-filters. The average effect of offering a no-opinion filter was again greatest for the least educated group ($d = .74$) and smallest for the most educated group ($d = .41$), with the medium-education group again falling in between ($d = .47$). The low-education group's effect was significantly larger than both the medium-education group's ($z = 1.88, p = .02$) and the high-education group's ($z = 2.39, p < .01$). The difference between the medium- and high-education groups was not significant ($z = 0.57, p = .28$).

Forbid/allow. In the forbid/allow experiments, the low- and medium-education groups' effect sizes ($d = .45$ and $.48$, respectively) were greater than the high-education group's ($d = .25$). The high-education group's effect was significantly smaller than that of the low-education group ($z = 3.08, p < .01$) and that of the medium-education group ($z = 3.89, p < .01$). The medium and low groups did not differ significantly ($z = 0.48, p = .32$).

Balance. Although Schuman and Presser (1981) concluded that unbalanced and formally balanced items yielded equivalent response distributions, a meta-analysis of their eight experiments indicated that there was in fact a significant difference between them ($d = .01, z = 2.57, p < .01$). As one would expect, formally balancing an unbalanced item decreased the number of respondents who answered "yes." However, this effect was quite small and essentially equal in the low-, medium-, and high-education groups ($d = .008, .004$, and $.007$).[7]

When gauged by comparisons with unbalanced and formally balanced items, the impact of salience enhancement was significantly smaller in the high-education group ($d = .22$) than in the low-education group ($d = .31, z = 1.87, p = .03$). The effect of salience enhancement in the medium-education group ($d = .28$) was not different from that in the low-education group ($z = 0.57, p = .29$), and the medium-education group's effect was marginally significantly different from the high-education group's ($z = 1.36, p = .09$).[8]

6. Most of the no-opinion filter experiments simply offered people the opportunity to say they had no opinion on the issue, whereas a few allowed respondents to say that they did not have sufficient interest in or information about the issue in order to offer an opinion one way or another. The relation of education to response effect magnitude was equally apparent when we analyzed the two sets of experiments separately.

7. The effect of formal balancing was significant for the full sample only for one of the eight experiments testing it, so no meta-analysis would have been possible using our usual approach. Therefore, we pooled all eight experiments to assess the moderating role of education in this case.

8. Schuman and Presser's (1981) experiments sometimes involved assigning respondents to receive one of three different question forms; such as unbalanced, formally balanced,

Question order. The question order experiments revealed a monotonic trend, with the effect size decreasing from the low-education group ($d = .42$) to the medium-education group ($d = .39$) to the high-education group ($d = .21$). The high group's effect was significantly smaller than the low group's ($z = 3.12$, $p < .01$) and the medium group's ($z = 3.21$, $p < .01$). The low and medium groups did not differ significantly ($z = .44$, $p = .33$).

Schuman and Presser (1981) explained the results of four of the question order experiments in terms of the norm of reciprocity. When we analyzed just these four experiments, the same basic results were obtained: the high-education group's effect size ($d = .15$) was significantly smaller than those of the medium-education group ($d = .65$, $z = 3.84$, $p = <.01$) and the low-education group ($d = .67$, $z = 3.34$, $p < .01$). The difference between the low- and medium-education groups was not significant ($z = 0.14$, $p = .44$).

In the remaining question order experiments, the same monotonic trend in effect sizes appeared ($d = .30$, .27, and .24 for the low-, medium-, and high-education groups, respectively). However, these effect sizes were not significantly different from one another. Thus, education did not moderate their magnitudes.[9]

Discussion

Summary. By employing meta-analytic techniques, we have generated more precise estimates of the relation between education and various response effects than an item-by-item approach permits. As a result, we found evidence indicating that education moderated the magnitudes of eight response effects: response order effects, acquiescence, middle alternative effects not involving status quo options, no-opinion effects for familiar issues, no-opinion effects for obscure issues, forbid/allow effects, balance effects involving salience enhancement, and question order effects due to the norm of reciprocity. In

---

and salience-enhanced. Such experiments therefore offer two opportunities to assess the impact of salience enhancement, by comparison with the unbalanced and formally balanced forms separately. However, these two comparisons would not be completely independent tests, because both would involve the same salience-enhanced condition. We chose to include only one comparison from each experiment in our meta-analysis. To be conservative, we dropped comparisons with unbalanced questions and included comparisons with formally balanced questions.

9. Of these nine items, the two abortion items were classified as part-whole contrast effects by Schuman and Presser (1981, p. 36). The general-marital happiness item and the occupational discrimination items were classified as part-whole consistency effects (Schuman and Presser, 1981, p. 42). The other five experiments were not classified at all (Schuman and Presser 1981, p. 52). We meta-analyzed the experiments in these three categories separately and found no evidence of moderation by education in any case.

each of these cases, the response effect was significantly greater in the low-education group than in the high-education group. Our evidence regarding acquiescence, no-opinion effects for obscure issues, forbid/allow effects, and question order effects due to the norm of reciprocity reinforces the validity of Schuman and Presser's (1981) conclusion that education moderated these effects. And our findings that education did not moderate other sorts of question order effects or the effect of formal balancing reinforces Schuman and Presser's (1981) view of these effects. But because education was found to moderate many other response effects, this body of evidence seems to contradict Schuman and Presser's (1981) final, general conclusion that education is not a "pervasive and systematic" moderator of response effects (p. 303). In fact, comparing the third and fifth columns of table 1 indicates that moderation by education appears to be the rule rather than the exception.

In considering our results, one might wonder whether the effects of education we have uncovered are large or small. In fact, meta-analysis experts argue strongly that labeling effect sizes in this way is unwise, because no such labels can be justified. For example, Glass, McGaw, and Smith (1981, p. 104) said, "There is no wisdom whatsoever in attempting to associate regions of the effect-size metric with descriptive adjectives such as 'small,' 'moderate,' 'large,' and the like." To label an effect as large or small depends on the standard of comparison, and selection of such a standard is to a large degree arbitrary. Furthermore, Rosenthal and Rubin (1982) have shown that, using the Binomial Effect Size Display, effects that appear to be small using conventional statistics appear quite large when considered in a different light. We are therefore inclined to view our results simply as suggesting a real and robust effect of education that may well help us to understand the etiology of these response effects.

Of course, the magnitude of education's impact did not exceed sampling error in the vast majority of the individual experiments we examined. So by this standard, the moderating role of education might be considered weak relative to the many social science effects that are statistically significant in individual studies. But this sort of interpretive approach may well be misleading. An independent variable may have sizable yet nonsignificant impact on a dependent variable if the latter is influenced by many other independent variables that are not included in the analysis and create what a statistical formula relegates to "error." And Krosnick (1991) has offered a theory proposing that some of the response effects examined here are indeed multiply determined by a wide range of factors. Consequently, we are reluctant to conclude that education's impact is small simply because its effect was often not statistically significant in individual studies.

Item-by-item variation. Trends toward stronger response effects among less educated respondents are apparent in majorities of the experiments yielding significant or marginally significant overall response effects (see the appendix). Specifically, the low-education group's response effect was larger than the high-education group's for all four of the norm of reciprocity experiments (100 percent), all nine of the middle alternative effect experiments not involving status quo options (100 percent), 10 of the 11 acquiescence experiments (91 percent), 20 of the 22 no-opinion effect experiments (91 percent), four of the five forbid/allow experiments (80 percent), five of the seven response order experiments (71 percent), and eight of the 12 salience-enhancement balance experiments (67 percent). So it does not appear that the significant meta-analytic results are generally the result of strong effects in minorities of experiments.

However, close inspection of the appendix suggests that education may have moderated effect size in some content domains but not others. For example, the two cases where education failed to moderate response order effects both involved the oil supply question. Similarly, the three cases in which education sizably moderated the forbid/allow effect all involved the item about speeches against democracy; no such education effects appeared for the forbid/allow items on other topics. These sorts of patterns may reveal useful information about the boundary conditions of education's effect, so it seems worthwhile for future studies to address whether the moderating impact of education is truly restricted to only some content domains or whether this is only illusory in the current data set.

In order to be more confident that variation in education's apparent impact across topics is not due simply to chance alone, we would like to formally assess its robustness. But unfortunately, we know of no statistical test that would allow us to assess whether the moderating effect of education was homogeneous or heterogeneous across a set of items. We are therefore hesitant to conclude too forcefully at this time that item content can limit education's effect. It would certainly be undesirable if meta-analysis were to lead us to ignore real variation in education's effect across content domains, but it would also be unproductive to take too seriously such variation in the absence of a sensible theoretical interpretation of it and evidence that it is indeed robust.

The medium-education group. The pattern of effect sizes for the medium-education group varied depending on the type of experiment involved. For all but one of the response effects that were moderated by education, the medium-education group's effect was significantly different from that of either the high-education group or the low-education group. In the response order experiments, acquiescence experiments, and the no-opinion experiments involving obscure issues, the effect sizes for the medium-education group were equivalent to those for the high-education group and were significantly smaller than those for the low-education group. But in the case of the middle alternative experiments not involving status quo options, no-opinion experiments involving familiar issues, the forbid/allow experiments, the salience-enhancement balance experiments, and the norm of reciprocity experiments, the effect sizes for the medium group were significantly or marginally significantly greater than those for the high group.

The rationales Schuman and Presser (1981) offered for why education might moderate response effect magnitude are certainly consistent with the general finding that various response effects are more prevalent among less educated respondents. But these investigators' general speculations cannot explain the different patterns of nonlinearity we observed in this relation. Certainly, it is possible that this variation is simply random noise; unfortunately, we know of no statistical test to evaluate whether these differences in the behavior of the medium-education group are statistically robust, so we should be cautious in taking them too seriously. Nonetheless, it seems worthwhile to speculate about what the observed pattern might mean, especially because it is understandable in light of Krosnick's (1991) recently articulated theory of survey satisficing.

According to this theory, optimizing entails carefully interpreting the meaning of each question, performing a comprehensive memory search to retrieve all relevant information, integrating this information carefully into a summary judgment, and reporting that judgment precisely by translating it onto the offered response choices. Although some respondents are undoubtedly motivated to optimize consistently in a survey, many may not be and may instead choose to satisfice: providing answers that will appear to be satisfactory or acceptable to the interviewer without having to fully execute all the steps of optimizing.

This can be done in two ways, one relatively subtle (weak satisficing) and the other more dramatic (strong satisficing). Weak satisficing occurs when respondents are less thorough in comprehension, retrieval, judgment, and/or response selection. Strong satisficing entails skipping the retrieval and judgment steps altogether, superficially interpreting a question, and selecting what will appear to be a reasonable answer to the interviewer and the researcher. When doing the latter, a respondent may look to the wording of the question for a cue, pointing to a response easily selected and easily defended if necessary.

The potential value of this theory is the outlining of factors likely to promote satisficing, one of which is the respondent's level of cognitive skills, the ensemble of talents a person may have to perform the cognitive operations necessitated by optimizing (i.e., linguistic interpretation, retrieval, integration, and expression of abstract ideas). The more such ability a person has, the easier it is to optimize, and the more likely he or she should be to do so. Educational attainment is very strongly correlated with direct assessments of cognitive skills (Ceci 1991), so respondents with relatively little education should be especially susceptible to satisficing-induced response effects.

Whereas highly skilled respondents seem unlikely to satisfice at all, and respondents with the least skills are probably consistently disposed to satisfice, the behavior of moderately skilled respondents may be contingent on the particular type of question being asked. When a question offers an easy opportunity to satisfice (i.e., a no-opinion option), respondents with moderate skills may recognize it as such and may be drawn to it, thus instigating strong satisficing, perhaps even at the same rate as among the least skilled respondents. But when a question does not offer an obvious satisficing answer choice (as in, e.g., a response order experiment), respondents with moderate skills may choose to optimize. Consequently, moderately skilled respondents may be no more likely to manifest weak satisficing than are highly skilled respondents.

Krosnick (1991) proposed that response order and acquiescence effects are likely to be the result of weak satisficing, and the behavior of the medium-education group here is consistent with this reasoning. That is, these individuals' effect sizes resembled those of the high-education group. And for the no-opinion filter effects involving familiar issues, the medium-education group's effect size was significantly stronger than that of the high-education group. This is consistent with Krosnick's presumption that the impact of such filters is due at least partly to strong satisficing. Likewise, the impact of offering non–status quo middle alternatives was greater among the medium-education groups than among the high-education group, which is consistent with the notion that such alternatives instigate strong satisficing.

In the no-opinion filter experiments involving obscure issues, the medium-education group's effect size resembled that of the high-education group. Like highly educated respondents, medium-education respondents may easily recognize that they have not heard of the issue, so a "don't know" response is the most appropriate one to give, regardless of whether it is offered explicitly or not.

We look forward to future research testing these many speculations about weak and strong satisficing and exploring the viability of this distinction more generally. In the meantime, this perspective offers a parsimonious account of some of the effects observed here.

Status quo middle alternatives. Although Krosnick (1991) suggested that status quo response options may be cues encouraging strong satisficing, the status quo option effects examined here were no greater among less educated respondents. This challenges the assertion that their impact is due to satisficing. It is not clear why the other middle alternative items manifested the education effect while it did not appear for the status quo items, and we look forward to future studies exploring this puzzle as well. In the meantime, our results suggest that analysts make a distinction between items offering prospective status quo middle alternatives and ones offering other sorts of middle alternatives.

Forbid/allow. Three of the response effects we examined are ones about which Krosnick's (1991) satisficing theory was mute (i.e., forbid/allow, balance, and question order), and statistically reliable education effects did occur in all three cases. As we mentioned above, education seemed to regulate the forbid/allow effect only for the item about speeches against democracy and not for items on other topics. As Schuman and Presser (1981) suggested, linguistic confusion caused by a sort of double negative ("*prohibit* public speeches *against* democracy") may have led some respondents to answer the democracy item inaccurately, and least educated respondents were presumably most prone to this mistake. So education may not moderate the forbid/allow effect generally but instead may determine the likelihood of response effects due to double-negative-induced linguistic confusion.

Balance. Although we found no variation across education groups in terms of the impact of formally balancing items, the impact of salience enhancement did vary with education. By according greater prominence to one point of view than another, salience-enhanced items apparently increase endorsement of that viewpoint. Less educated respondents may be particularly inclined to select a response alternative if it is especially salient (perhaps due to satisficing), so salience enhancement may have the most impact on these individuals for this reason. If this impact is indeed due to satisficing, the resemblance of the medium-education group's effect size to that of the low-education group suggests that endorsing more extensively presented viewpoints may be a form of strong satisficing.

Question order. Our evidence that less educated respondents were more susceptible to question order effects involving the norm of reciprocity is also interpretable from the satisficing viewpoint. Respondents who optimize when answering a question on communist newspaper reporters, for example, should be likely to recognize on their own that the norm of reciprocity applies and has implications for U.S. reporters as well. So the second question in a pair, which would normally activate the norm, should have little or no effect, because these individuals have already thought of and applied the norm when answering the first question. But individuals inclined to satisfice are presumably unlikely to think of the norm on their own when answering the first question. Consequently, activation of the norm by the second question should be especially pronounced among them. Activation of the norm is probably a powerful cue, suggesting a response to the second question that is easy to defend, thereby perhaps instigating a strong form of satisficing. Our results consistent with this logic (i.e., the medium-education group resembled the low-education group) encourage further tests of whether the norm of reciprocity effects are most likely to occur under all the conditions thought to foster satisficing.

Implications for survey design. Of course, our results do not provide any indication of the frequency with which response effects occur in surveys. But when some such effects occur, they are apparently more likely among respondents with lower levels of education. One could view this finding as auguring bleak prospects for the hope that surveys might be designed in ways that curb such response effects. Educational attainment is essentially an immutable attribute of individuals, so its impact in enhancing the likelihood of satisficing may seem unavoidable, bound to lower the quality of data obtained from some respondents.

However, the satisficing perspective presumes that steps can be taken in survey design to significantly reduce the likelihood of satisficing even among the least educated respondents. Specifically, Krosnick (1991) suggested that minimizing the difficulty level of survey questions and maximizing respondent motivation to expend efforts to answer survey questions may dramatically enhance optimizing. Using simple vocabulary, minimizing distractions, providing instructions requesting accurate data (see, e.g., Oksenberg, Vinokur, and Cannell 1979a, 1979b), holding respondents accountable for their answers (see, e.g., Tetlock 1983a; Tetlock and Kim 1987), and many other such tactics may all be effective in this regard. Consequently, we look forward to future research exploring the interactive and compensatory relations among various of such factors in moderating the magnitude of response effects and thereby affecting the reliability and validity of survey responses.

Conclusion

In conclusion, it is useful to note that our findings nicely complement those reported by Krosnick and Schuman (1988) a few years ago. Those investigators tested the widely held assumption that response effects in attitude measurement are most likely to appear among individuals whose attitudes are weak (see, e.g., Cantril 1944; Converse 1974). According to this view, cues in questions or psychological forces exerted by their wording, format, or order can most easily push around individuals who have only weak internal cues indicating their attitudes.

Early studies of this hypothesis had been conflicting and therefore inconclusive (see, e.g., Sudman and Bradburn 1974; Sudman and Swensen 1985), so Krosnick and Schuman's (1988) meta-analysis of Schuman and Presser's (1981) 27 relevant experiments was useful in providing a more definitive test of the notion. Surprisingly, Krosnick and Schuman (1988) found that measures of attitude importance, intensity, and certainty did not identify respondents especially susceptible to response effects. Thus, their study left this literature confronted with the mystery of what individual difference variable(s), if any, could account for susceptibility to these effects. The present investigation, of course, has identified just such a variable (i.e., education) and thereby provides a justification for further careful study of the precise psychological mechanisms responsible for its moderating effects.

Appendix

Table A1. Summary of the Results of Individual Studies

Experiment	Overall Response Effect (%)[a]	Significance of Overall Effect	Response Effect[b] Low Education	Medium Education	High Education	Three-Way Interaction χ^2	p	N
Response order:								
Oil companies:								
April 1979	7.64	$p < .10$	31.91*	2.10	3.29	6.95	.03	655
Oil supply:								
January 1979	14.03	$p < .001$	17.33⁺	6.80	19.80*	1.94	.16	555
April 1979	8.76	$p < .05$	8.29	1.48	16.40*	2.82	.24	656
Divorce:								
April 1979	12.39	$p < .01$	20.91⁺	4.20	16.56*	4.55	.34	676
September 1979	10.77	$p < .05$	16.93	4.44	12.86	2.99	.56	577
Adequate housing:								
August 1979	14.39	$p < .001$	19.50*	13.02*	13.63*	.26	.88	619
September 1979	7.39	$p < .10$	6.67	11.29⁺	4.87	.59	.74	576
Oil companies:								
January 1979	6.09	n.s.	2.07	1.30	12.82*	1.68	.43	547
Vietnam aid:								
April 1979	3.57	n.s.	−2.19	3.31	3.20	3.93	.41	608
Unions:								
June 1979	2.56	n.s.	14.77	−5.96	−4.96	4.61	.09	622
Individuals/social conditions:								
December 1978	1.33	n.s.	15.28	4.15	−6.14	2.94	.23	466
Open housing:								
November 1979	2.25	n.s.	−3.70	9.88	−2.49	2.64	.27	690

Table A1 (*Continued*)

Experiment	Overall Response Effect (%)[a]	Significance of Overall Effect	Response Effect[b] Low Education	Medium Education	High Education	Three-Way Interaction χ^2	p	N
Divorce:								
June 1979	.27	n.s.	−.89	5.85	−2.42	.85	.65	545
Individuals/social conditions:[c]								
February 1980	.38	n.s.	13.15	−5.5	.20	3.86	.14	927
Work values:[d]								
February 1977		n.s.				27.96	.67	582
August 1977		n.s.				4.01	.85	543
Most important problem:[e]								
February 1977		n.s.				29.12	.61	581
Acquiescence:								
Women in politics:								
Fall 1974	6.30	$p < .05$	15.25*	1.17	4.49	4.40	.11	1,439
National Opinion Research Center (NORC) 1974	14.00	$p < .001$	23.28*	6.99	10.61*	6.68	.04	1,409
February 1976	11.79	$p < .001$	14.81*	11.25*	9.12*	.33	.85	1,223
Spring 1976	9.75	$p < .001$	12.89*	10.67*	6.92*	1.30	.52	2,913
June 1980	6.61	$p < .10$	25.23*	2.68	4.33	4.36	.11	625
Individuals/social conditions:								
Fall 1974	13.13	$p < .001$	12.35*	17.44*	11.16*	.67	.72	916
Russian leaders:								
Fall 1974	9.76	$p < .01$	15.66*	5.10	6.31	1.63	.44	818
Politics:								
Detroit Area Study (DAS) 1976	5.01	$p < .10$	1.43	5.23	6.19	.17	.92	1,116
Voting:								
DAS 1976	11.96	$p < .001$	12.19*	15.86*	9.06+	.68	.71	1,099
Jobs:								
DAS 1976	9.79	$p < .001$	11.12*	9.50+	5.96	.77	.68	1,117

Adequate housing:								
DAS 1976	23.93	$p < .001$	32.41*	15.95*	22.63*	6.33	.04	1,085
Individuals/social conditions:								
June 1980	2.12	n.s.	6.33	−1.04	.39	1.36	.51	616
Medical care:								
DAS 1976	2.85	n.s.	2.50	3.52	1.31	.25	.88	1,098
Russian leaders:								
August 1979	2.93	n.s.	−4.77	2.83	5.34	.71	.70	565
Middle alternative—status quo:								
Marijuana:								
Fall 1974	12.56	$p < .001$	5.73*	17.09*	15.37*	3.17	.20	1,430
February 1976	13.44	$p < .001$	8.84*	12.86*	16.36*	2.11	.35	1,200
August 1977	19.73	$p < .001$	16.95*	16.62*	23.24*	.68	.71	1,119
Local education:								
February 1976	16.32	$p < .001$	12.22*	22.02*	13.12*	2.18	.33	1,086
Divorce:								
NORC 1978	13.49	$p < .001$	16.42*	15.84*	7.98*	3.13	.20	2,171
September 1979	25.64	$p < .001$	16.41*	27.65*	28.75*	.02	.99	568
Middle alternative—others:								
Aid to Vietnam:								
Fall 1974	11.89	$p < .001$	12.53*	17.11*	4.75	2.52	.28	1,309
February 1975[f]	9.97	$p < .001$	18.19*	9.51*	6.00	2.40	.30	1,237
February 1978	14.20	$p < .001$	24.24*	25.86*	7.55*	5.01	.08	1,018
August 1978	13.21	$p < .001$	13.92*	16.99*	9.26*	.76	.68	964
Fall 1978	14.46	$p < .001$	23.33*	15.73*	9.82*	1.22	.54	1,240
Liberal/conservative:								
Fall 1974	39.27	$p < .001$	41.54*	43.70*	33.87*	.15	.92	946
Spring 1976	40.88	$p < .001$	46.11*	47.66*	31.79*	6.09	.04	2,778
February 1978	22.98	$p < .001$	26.71*	35.47*	18.16*	10.07	.01	1,086
Fall 1978[g]	46.23	$p < .001$	62.12*	54.77*	34.50*	11.89	.00	939
No opinion filter—familiar issues, full filters:								
Russian leaders:								
Fall 1974	22.46	$p < .001$	31.35*	24.40*	14.04*	.25	.88	1,007
May 1978	33.34	$p < .001$	37.54*	44.53*	22.68*	5.45	.06	1,212
November 1978	27.50	$p < .001$	27.67*	29.31*	25.52*	1.72	.42	1,535

Table A1 (*Continued*)

Experiment	Overall Response Effect (%)[a]	Significance of Overall Effect	Response Effect[b] Low Education	Medium Education	High Education	Three-Way Interaction χ^2	p	N
Middle East:								
Fall 1974	21.90	$p < .001$	31.60*	26.60*	10.74*	3.39	.18	1,003
February 1975	28.97	$p < .001$	17.00*	48.82*	19.78*	23.19	.00	1,361
Spring 1976	32.07	$p < .001$	36.69*	33.74*	26.82*	4.51	.10	3,066
May 1978	30.55	$p < .001$	50.28*	34.30*	17.22*	9.46	.01	1,185
Portugal:								
Fall 1974	24.86	$p < .001$	24.63*	22.34*	27.13*	2.40	.30	1,007
Government power:								
DAS 1976	15.00	$p < .001$	18.23*	19.79*	6.64+	2.42	.30	1,113
February 1976	20.99	$p < .001$	20.76*	24.69*	19.69*	7.14	.02	1,198
March 1978	22.85	$p < .001$	26.20*	22.84*	19.71*	.22	.89	707
No opinion filter—familiar issues, quasi filters:								
Government leaders:								
May 1978	16.31	$p < .001$	22.37*	20.19*	10.28*	3.71	.16	1,162
Government crooks:								
May 1978	14.83	$p < .001$	22.02*	18.06*	8.72*	3.44	.18	1,203
Fall 1978	12.62	$p < .001$	17.04*	14.82*	8.84*	1.77	.41	1,477
Communist book:								
February 1977	23.63	$p < .001$	40.60*	28.62*	13.31*	.42	.81	1,103
March 1978	22.69	$p < .001$	39.96*	18.97*	15.01*	2.39	.30	705
May 1978	19.66	$p < .001$	31.91*	24.09*	10.30*	.71	.70	1,187
Courts:								
NORC 1974	23.44	$p < .001$	29.79*	14.48*	25.47*	9.63	.01	1,343
DAS 1976	10.54	$p < .001$	9.80*	13.04*	8.23*	1.28	.53	1,070
Liberal/conservative:								
NORC 1978	12.15	$p < .001$	17.64*	13.82*	5.14*	1.21	.54	1,503

No opinion filter—obscure issues:								
Agricultural trade act:								
Fall 1978	20.51	$p < .001$	33.75*	17.92*	16.14*	4.74	.09	1,164
Monetary Control Bill:								
April 1979	19.71	$p < .001$	27.63*	21.04*	14.84*	1.06	.59	692
Forbid/allow:								
Democracy forbid/allow:								
Fall 1974	15.61	$p < .001$	25.56*	18.75*	5.94	5.75	.05	1,422
February 1976	25.04	$p < .001$	30.72*	34.35*	11.83*	5.26	.07	1,177
Spring 1976	26.34	$p < .001$	33.44*	33.42*	16.09*	8.36	.02	2,799
Communism forbid/allow:								
DAS 1976	18.01	$p < .001$	20.00*	16.13*	20.10*	1.07	.58	1,050
X-rated movies:								
August 1977	4.81	$p < .10$	.45	11.65*	.32	2.94	.23	1,107
Cigarette ads:								
September 1979	4.98	n.s.	12.50	3.75	2.34	1.16	.55	871
Balance:								
Unbalanced vs. formal balance:								
Fuel shortage:								
Fall 1974	8.87	$p < .001$	7.60	11.50*	8.02^{+}	.46	.79	998
February 1977	1.15	n.s.	4.03	.07	.52	.26	.88	1,106
August 1978	4.54	n.s.	7.32	−3.02	9.34	1.74	.42	538
November 1978	4.61	n.s.	1.46	2.29	8.05	.99	.60	976
Unions:								
Fall 1974	2.27	n.s.	5.00	3.94	−.90	.62	.73	974
Abortion:								
February 1975	2.08	n.s.	6.70	.13	3.56	.59	.74	865
Gun permits:								
February 1975	−.60	n.s.	−.99	−.77	.10	.02	.99	896
Fall 1978	.58	n.s.	2.07	−.32	−.96	.09	.95	970

Table AI (*Continued*)

Experiment	Overall Response Effect (%)[a]	Significance of Overall Effect	Response Effect[b] Low Education	Medium Education	High Education	Three-Way Interaction χ^2	p	N
Unbalanced vs. salience-enhanced:								
Fuel shortage:								
Fall 1974	12.14	$p < .001$	18.14*	14.98*	4.88	3.22	.20	983
Fall 1978[h]	17.08	$p < .001$	13.22*	15.59*	20.67*	1.50	.47	980
Unions:								
Fall 1974[h]	10.78	$p < .001$	24.47*	9.26[+]	4.37	3.73	.15	964
February 1975	5.77	n.s.	5.49	7.76[+]	3.01	.23	.89	1,334
Abortion:								
February 1975	6.36	$p < .05$	4.82	10.13[+]	5.97	.48	.79	859
Gun permits:								
Fall 1978[h]	8.37	$p < .01$	21.69*	7.03	2.36	4.90	.08	959
February 1975	3.73	n.s.	6.41	6.97	−.64	1.22	.54	883
September 1979	5.87	n.s.	6.71	15.85*	−3.40	4.39	.11	569
Formal balance vs. salience-enhanced:								
Fuel shortage:								
Fall 1978	12.47	$p < .001$	11.76*	13.30*	12.62*	.12	.94	966
September 1980	29.21	$p < .001$	30.50*	25.82*	31.70*	.55	.76	623
Fall 1974	3.27	n.s.	10.54*	3.48	−3.14	3.38	.18	969
Unions:								
Fall 1974	8.51	$p < .005$	19.47*	5.32	5.27	2.56	.27	946
September 1980	11.44	$p < .005$	12.64	17.81*	7.24	.65	.72	606
Abortion:								
February 1975	4.28	n.s.	−1.88	10.00[+]	2.41	2.04	.36	850

Gun permits:								
February 1976	6.37	$p < .01$	6.98	9.83*	3.46	1.09	.58	1,200
Fall 1978	7.79	$p < .05$	19.62*	7.35	1.40	4.50	.10	943
February 1975	4.33	n.s.	7.40	7.74	−.74	1.59	.45	875
Vietnam:								
Fall 1974[i]	15.14	$p < .001$	19.38*	13.33*	12.92*	1.01	.63	1,324
February 1976	12.22	$p < .001$	10.13*	17.54*	9.24*	2.11	.35	1,124
Fall 1978	18.50	$p < .001$	24.62*	13.97*	18.91*	2.45	.29	1,313
Communist book:								
February 1976[j]	5.15	$p < .05$	.90	9.29*	1.46	2.05	.36	1,216
Fall 1974[k]	1.99	n.s.	−15.06*	6.16	8.33*	13.17	.00	958
Fall 1974[l]	4.72	n.s.	−4.68	8.08⁺	9.52*	6.81	.03	973
Atheist:								
Fall 1974	.60	n.s.	−.78	.47	.90	.08	.96	1,008
Question order:								
Norm of reciprocity:								
Foreign reporter (Communist):								
May 1980	19.35	$p < .001$	33.97*	26.71*	4.17	4.82	.09	350
July 1980	23.35	$p < .001$	23.59⁺	39.25*	8.73	2.60	.27	310
Foreign reporter (American):								
May 1980	16.94	$p < .001$	30.32*	20.02*	5.86	1.81	.40	342
July 1980	20.23	$p < .001$	38.69*	32.31*	3.99	4.36	.11	308
Other question order experiments:								
Abortion (general):								
June 1979	13.38	$p < .001$	13.04	13.45*	13.89*	.02	.99	599
August 1979	18.56	$p < .001$	27.18*	13.21*	18.31*	1.99	.37	608
Occupational discrimination (general):								
February 1980	5.42	$p < .05$	10.47⁺	9.06*	.45	4.02	.13	451
Marital happiness (general):								
August 1979	13.69	$p < .01$	5.00	19.66*	11.53	4.62	.33	377
Public officials care:								
DAS 1971	8.55	$p < .001$	3.09*	9.29*	14.27*	7.64	.26	1,839

Table A1 (*Continued*)

Experiment	Overall Response Effect (%)[a]	Significance of Overall Effect	Response Effect[b]			Three-Way Interaction		
			Low Education	Medium Education	High Education	χ^2	p	N
Say in government:								
DAS 1971	7.02	$p < .001$	4.21*	8.21*	8.87*	9.80	.13	1,850
Politics:								
DAS 1971	5.24	$p < .001$	.04+	.14	8.57*	3.17	.79	1,851
Science:								
DAS 1971	.43	$p < .01$	1.90	1.90	−.23	1.11	.98	1,835
Public officials don't care:								
DAS 1971	4.16	$p < .001$	7.21*	2.21	3.04	6.82	.33	1,839
Doctor/lawyer (lawyer item):								
Fall 1974	3.82	n.s.	4.02	3.12	3.19	.11	.95	1,348
Doctor/lawyer (doctor item):								
Fall 1974	4.24	n.s.	8.46+	4.29	1.19	1.08	.58	1,366
Woman/Jewish president (woman item):								
February 1975	2.90	n.s.	2.63	5.29+	.33	1.32	.52	1,332
Woman/Jewish president (Jewish item):								
February 1975	1.69	n.s.	6.85	1.28	.31	1.01	.60	1,334
Abortion (specific):								
August 1979	1.96	n.s.	6.66	3.34	1.25	1.11	.57	617
June 1979	1.04	n.s.	1.12	.78	1.09	.01	.99	610
Occupational discrimination (specific item):								
February 1980	.38	n.s.	5.27	1.01	−1.90	.52	.77	437
Marital happiness (specific item):								
August 1979	7.29	n.s.	−.19	8.55	8.92	2.01	.73	376

Vote:								
DAS 1971	4.10	n.s.	.76	−1.55	1.17	5.12	.53	1,866
Voting only way:								
DAS 1971	.05	n.s.	3.26	0.92	−4.35	2.53	.87	1,849

[a] This column reports the difference in the percentage of respondents giving the same answer on two different versions of a question. For example, for the oil supply response order item, the overall response effect is the difference between the percentage of respondents saying "plenty of oil" on the form where it is offered as the first of two options and the form where it is offered as the second of two options. Thus, this figure represents the magnitude of the effect of the question manipulation in percentage terms. The result in this column does not always agree precisely with that reported by Schuman and Presser (1981) because we confined our analyses to those respondents whose education attainment was ascertained.

[b] The three subcolumns under the heading "Response Effect" report the difference in percentage of respondents giving the same response on two different versions of a question. This difference was computed separately for subsamples of respondents in the low-, medium-, and high-education groups. The levels of significance attached to these figures test whether the question manipulation effect was statistically significant within each education group.

[c] Schuman and Presser (1981) report the results of this response order experiment in chap. 8 of their book in n. 4. This experiment involved asking some respondents, "Please tell me whether you agree or disagree with this statement: Individuals are . . ."; other respondents were asked, "Please tell me whether you disagree or agree with this statement: Individuals are. . . ."

[d] These experiments involved presenting five substantive response alternatives in five different orders, and three alternatives in three orders, respectively. Consequently there is no simple way to summarize the sizes of the effects. The effect size columns in this row are therefore left blank. The experiment involving three different orders of three response alternatives has been reported by Schuman and Presser (1981, p. 103).

[e] This experiment also involved presenting five substantive response alternatives in five different orders, so the effect size columns in this row are also left blank.

[f] This middle alternative experiment was conducted by Schuman and Presser (1981) but was not described in their book.

[g] In the liberal/conservative middle alternative experiment conducted in fall 1978, respondents were asked one of three question versions, involving either five response options (liberal, somewhat liberal, middle-of-the-road, somewhat conservative, and conservative), three options (liberal, middle-of-the-road, conservative), or two options (liberal, conservative). In order to focus solely on the impact of offering the middle alternative, we did not use the data from the five-choice form and simply compared the three-choice form to the two-choice form.

[h] These comparisons were not included in the meta-analysis because they involved comparisons that were not independent of others that were included, and these latter comparisons were more conservative.

[i] Schuman and Presser (1981) treated their experiments asking about sending troops if "another Vietnam" were to occur as examining tone of wording effects. However, the manipulation simply involved adding the phrase "to stop a Communist takeover" to a formally balanced question. Because this seems comparable to the balance experiments, where an argument in favor of one alternative was added to an unbalanced or formally balanced question, we treated these experiments as testing for balance effects rather than tone of wording.

[j] This involves a comparison of the "Stouffer" vs. the "liberal" versions of the Communist book question. The difference between these two versions is simply the addition of an argument in favor of one side: "Somebody else in your community says this is a free country and it should be allowed to remain." We therefore treated this comparison as assessing the impact of the argument, under the rubric of balance experiments.

[k] This involves the same comparison as is described in n. i above.

[l] This involves a comparison of the "neutral" form and the "Stouffer" form of the Communist book question. Again, the difference here is the addition of an argument in favor of one side: "Somebody in your community suggests the book should be removed from the library." We therefore treated this comparison as assessing the impact of the argument, under the rubric of balance experiments.

$^{+}$ $p < .10$.

* $p < .05$.

References

Cantril, Hadley. 1944. *Gauging Public Opinion.* Princeton, NJ: Princeton University Press.

Ceci, Steven J. 1991. "How Much Does Schooling Influence General Intelligence and Its Cognitive Components? A Reassessment of the Evidence." *Developmental Psychology* 27:703–22.

Converse, Philip E. 1974. "Comment: The Status of Nonattitudes." *American Political Science Review* 68:650–60.

Glass, Gene V., Barry McGaw, and Mary Lee Smith. 1981. *Meta-analysis in Social Research.* Beverly Hills, CA: Sage.

Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5:213–36.

Krosnick, Jon A., and Howard Schuman. 1988. "Attitude Intensity, Importance, and Certainty and Susceptibility to Response Effects." *Journal of Personality and Social Psychology* 54:940–52.

Mullen, Brian. 1989. *Advanced BASIC Meta-Analysis.* Hillsdale, NJ: Erlbaum.

Oksenberg, Lois, Amiram Vinokur, and Charles F. Cannell. 1979a. "Effects of Commitment to Being a Good Respondent on Interview Performance." In *Experiments in Interviewing Techniques,* ed. Charles F. Cannell, Lois Oksenberg, and Jean M. Converse. Ann Arbor: Survey Research Center, University of Michigan.

———. 1979b. "The Effects of Instructions, Commitment, and Feedback on Reporting in Personal Interviews." In *Experiments in Interviewing Techniques,* ed. Charles F. Cannell, Lois Oksenberg, and Jean M. Converse. Ann Arbor: Survey Research Center, University of Michigan.

Oliver, L. W. 1987. "Research Integration for Psychologists: An Overview of Approaches." *Journal of Applied Social Psychology* 17:860–74.

Rosenthal, Robert. 1988. *Meta-Analytic Procedures for Social Research.* Beverly Hills, CA: Sage.

Rosenthal, Robert, and D. B. Rubin. 1982. "A Simple, General Purpose Display of Magnitude of Experimental Effect." *Journal of Educational Psychology* 74:166–69.

Schuman, Howard, and Graham Kalton. 1985. "Survey Methods." In *The Handbook of Social Psychology,* ed. Gardner Lindzey and Eliot Aronson. 3d ed. New York: Random House.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in the Attitude Surveys: Experiments on Question Form, Wording, and Context.* New York: Harcourt Brace Jovanovich.

Sudman, Seymour, and Norman M. Bradburn. 1974. *Response Effects in Surveys.* Chicago: Aldine.

Sudman, Seymour, and K. Swensen. 1985. "Measuring the Effects of Attitude Crystallization on Response Effects." Paper presented at the annual meeting of the American Association for Public Opinion Research, McAfee, NJ.

Tetlock, Philip. 1983. "Accountability and Complexity of Thought." *Journal of Personality and Social Psychology* 45:74–83.

Tetlock, Philip, and Jae I. Kim. 1987. "Accountability and Judgment Processes in a Personality Prediction Task." *Journal of Personality and Social Psychology* 52: 700–709.

Tourangeau, Roger, and Kenneth A. Rasinski. 1988. "Cognitive Processes Underlying Context Effects in Attitude Measurement." *Psychological Bulletin* 103:299–314.